PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 9/04, 15/53

A1

(11) International Publication Number: WO 95/30743

(43) International Publication Date: 16 November 1995 (16.11.95)

(21) International Application Number: PCT/DK95/00180

(22) International Filing Date: 4 May 1995 (04.05.95)

(30) Priority Data:
0509/94 4 May 1994 (04.05.94) DK

(71) Applicant *(for all designated States except US)*: NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsvaerd (DK).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: VON DER OSTEN, Claus [DK/DK]; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd (DK). FREDHOLM, Henrik [DK/DK]; Novo Nordisk A/S, Novo Allé, DK-2880 Bagsvaerd (DK).

(74) Common Representative: NOVO NORDISK A/S; Corporate Patents, Novo Allé, DK-2880 Bagsvaerd (DK).

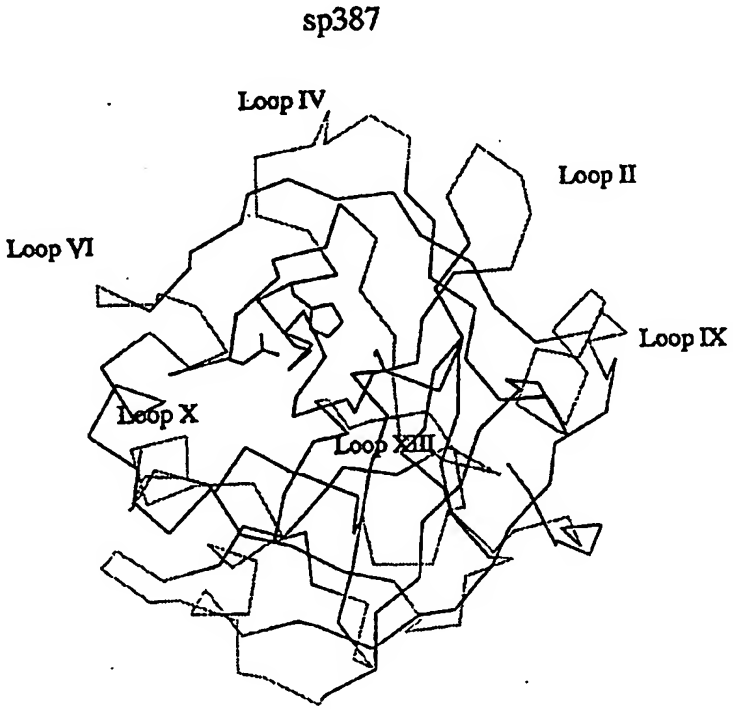
(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PROTEASE VARIANTS

(57) Abstract

The present invention relates to novel trypsin-like protease variants with improved properties, DNA constructs coding for the expression of said variants, host cells capable of expressing the variants from the DNA constructs, as well as a method of producing the variants by cultivating said host cells. The variants may advantageously be used as constituents in detergent compositions and additives.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

PROTEASE VARIANTS

FIELD OF THE INVENTION

The present invention relates to novel trypsin-like protease variants with improved properties, DNA constructs coding for the expression of said variants, host cells capable of expressing the variants from the DNA constructs, as well as a method of producing the variants by cultivating said host cells. The variants may advantageously be used as constituents in detergent compositions and additives.

BACKGROUND OF THE INVENTION

Trypsin-like proteases, i.e. serine-proteases that in structure are similar to trypsin, have been extensively described in the literature (A J. Greer, "Comparative modelling methods - application to the family of mammalian serine proteases", Proteins, Vol. 7, p. 317-334, 1990; M. A. Phillips & R. J. Fletterick, "Proteases", Curr. Opin. Struct. Biol., Vol. 2, p. 713-720, 1992).

The trypsin-like proteases can be divided into different families bases on their structure (A. Sali & T. Blundell, "Definition of general topological equivalence in protein structures", J. Mol. Biol., 212, p. 403-428, 1990; J. P. Overington et al., "Environment-specific amino acid substitution tables: Tertiary templates and prediction of protein folds", Protein Science, Vol. 1, p. 216-226). One such family of closely structurally related trypsin-like proteases comprises eight mammalian and one bacterial trypsin-like proteases, i.e. the following nine proteases: Trypsin (Rattus rattus), PDB Code 1trm, S. Sprang et al., "The three-dimensional structure of Asn102 mutant of trypsin", Science, Vol. 235, p. 905, 1987; Trypsin (Bos taurus), PDB Code 2ptn, J. Walter et al., "On the disordered activation domain in trypsinogen", Acta Cryst., Vol. 38B, p. 1462, 1982; Tonin

(Rattus rattus), PDB Code 1ton, M. Fujinaga et al., "Rat submaxillary gland serine protease, Tonin", J. Mol. Biol., Vol. 195, p. 373, 1987; Kallikrein A (Sus scrofa), PDB Code 2pka, W. Bode et al., "Refined 2 Ångstroms X-ray crystal structure of porcine pancreatic Kallikrein A", J. Mol. Biol., Vol. 164, p. 273, 1983; Gamma-chymotrypsin (Bos taurus), PDB Code 2gch, G. H. Cohen et al., "Refined crystal structure of Gamma-chymotrypsin at 1.9 Ångstroms resolution", J. Mol. Biol., Vol. 148, p. 449, 1981; Pancreatic Elastase (Sus scrofa), PDB Code 3est, E. Meyer et al., PDB Code 3est, E. Meyer et al., "Structure of native porcine pancreatic elastase at 1.65 Ångstroms resolution'', Acta Cryst., Vol. 44B, p. 26, 1988; Neutrophil Elastase (Homo sapiens), PDB Code 1hne, M. A. Navia et al., "Structure of Human Neutrophil Elastase in complex with a peptide chloromethyl ketone inhibitor at 1.84 Ångstroms resolution", Proc. Nat. Acad. Sci. USA, Vol. 86, p. 7, 1989; Mast Cell Proteinase (Rattus rattus), PDB Code 3rp2, S. J. Remington et al., "The structure of Rat Mast Cell protease II at 1.9 Ångstoms resolution", Biochemistry, Vol. 27, p. 8097, 1988.

The PDB Codes refer to the structural data files deposited at the Brookhaven Protein Data Bank (F. C. Bernstein et al., "The Protein Data Bank: A computer based archival file for macromolecular structures", J. Mol. Biol., Vol. 112, p. 535-542, 1977.

WO 89/06270 discloses a trypsin-like protease obtained from a strain of the fungal species Fusarium oxysporum as well as a detergent composition comprising the enzyme. No information as to the amino acid sequence of the enzyme or the three-dimensional structure thereof have been published.

SUMMARY OF THE INVENTION

The trypsin-like F. oxysporum protease described in WO 89/06270 has now been cloned and sequenced and the three-dimensional structure thereof has been elucidated by X-ray crystallography. This trypsin-like protease has surprisingly

been found to belong to the above mentioned trypsin family comprising eight known mammalian and one bacterial trypsin-like protease. From a comparison with the structure of the trypsins belonging to this family it has surprisingly been shown that while the structures are very similar near the active and binding sites and in the core of the protein, there are considerable differences in other regions, particularly in the loops on the surface of the molecule. Furthermore, there is no evidence of any divalent cation binding sites in the F. oxysporum trypsin-like protease.

It is the object of the present invention to design novel variants of a trypsin-like Fusarium protease having improved properties as compared to those of the parent trypsin-like protease.

Accordingly, in a first aspect the present invention relates to a variant of a parent trypsin-like Fusarium protease, which

i) reacts with an antibody raised against or reactive with at least one epitope of the trypsin-like F. oxysporum protease comprising the amino acid sequence shown in the appended Sequence Listing ID No. 2,

ii) is at least 60% homologous with the amino acid sequence of the trypsin-like F. oxysporum protease shown in SEQ ID No. 2, and/or

iii) is encoded by a DNA sequence which hybridizes with an oligonucleotide probe hybridizing with a DNA sequence encoding the trypsin-like F. oxysporum protease having the amino acid sequence shown in SEQ ID No. 2.

In the present context, the term "trypsin-like Fusarium protease" is intended to indicate a trypsin-like protease derived from a fungus of the genus Fusarium, and in particular of the species F. oxysporum, or a functional analogue thereof.

The term "functional analogue" is intended to indicate a trypsin-like protease which is immunologically cross-reactive with the trypsin-like F. oxysporum protease described herein,

comprises an amino acid sequence which is more than 60% homologous with that of the trypsin-like F. oxysporum protease shown in SEQ ID No. 2, such as more than 70%, 80% or even 90% with said protease, is encoded by a DNA sequence hybridizing with an oligonucleotide probe which also hybridizes with a DNA sequence encoding the trypsin-like F. oxysporum protease the amino acid sequence of which is shown in SEQ ID No. 2, and/or has a three-dimensional structure having a core which is substantially similar to the core of the trypsin-like F. oxysporum protease and which preferably has one or more loop structures corresponding to loops II, IV, VI, IX, X and XIII of the trypsin-like F. oxysporum protease described herein. The term "corresponding" as used about the loop structures is intended to indicate an identity of at least 60% such as 70%, 80%, 90% or up to 100% with the corresponding F. oxysporum protease loop structure(s). The properties characterizing the functional analogue are intended to be understood in an analogous manner to properties i)-iii) listed above and further described below.

For ease of reference, the following disclosure is based on the trypsin-like protease derived from the species F. oxysporum (DSM 2672), the cDNA and amino acid sequences of which are apparent from SEQ ID Nos. 1 and 2, respectively. It will be understood, however, that also functional analogues of said protease as defined above, e.g. trypsin-like proteases derivable from other organisms such as microorganisms including bacterial and fungal strains, and in particular from other strains of Fusarium spp., may be modified in a manner similar to that described for the trypsin-like F. oxysporum protease described herein. Accordingly, variants of such functional analogous are intended to be considered to be within the scope of the present invention. Examples of other Fusarium strains, which have been found to produce trypsin-like proteases, are F. merismoides, F. redolens, F. sambucinum, F. solani and F. verticilloides.

The term "variant" is intended to indicate a polypeptide which is derived from a trypsin-like Fusarium protease as defined above and which has one or more of the properties i)-iii) which will be further discussed below. Typically, the variant differ from the trypsin-like protease by one or more amino acid residues, which, for instance, may have been added or deleted from either or both of the N-terminal or C-terminal end of the protease, inserted or deleted at one or more sites within the amino acid sequence of the protease, or substituted with one or more amino acid residues within, or at either or both ends of the amino acid sequence of the protease.

As stated above, the comparison of the three-dimensional structure of the trypsin-like Fusarium protease (shown in Fig. 1) with that of other known trypsins revealed a considerable difference in the surface structures of the proteases. It is contemplated that properties of other trypsin-like proteases may be improved when loop structures of the trypsin-like F. oxysporum protease disclosed herein are inserted in or substituted for loop structures of such trypsin-like proteases.

Accordingly, in a further aspect the present invention relates to a variant of a parent trypsin-like protease comprising at least one of the loop structures of the trypsin-like Fusarium protease.

In the present context the term "trypsin-like protease" is intended to indicate an enzyme having a three-dimensional structure similar to that of the class of trypsins listed in Table 2, below. It will be understood that trypsin as such is considered to be included within this definition.

The variants of the present invention are contemplated to have improved substrate specificities, catalytic rate, stability, especially towards the action of proteolytic enzymes and/or detergent ingredients, thermostability, storage stability, improved resistance towards peroxidase/pHBS inactivation, and/or improved wash performance.

The present invention also relates to a DNA construct comprising a DNA sequence encoding a trypsin-like protease variant as indicated above, a recombinant expression vector carrying said DNA construct, a cell transformed with the DNA construct or the expression vector, as well as a method of producing a trypsin-like protease variant of the invention by culturing or growing said cell under conditions conducive to the production of the variant, after which the variant is recovered from the culture.

The invention further relates to a enzyme granulate, a liquid enzyme composition or a protected enzyme preparation comprising a trypsin-like protease variant of the invention and suitable for the preparation of e.g. a detergent composition comprising a trypsin-like protease variant of the invention.

BRIEF DESCRIPTION OF DRAWINGS

The present invention is further illustrated by reference to the accompanying drawings, in which:

Fig. 1 shows the three-dimensional structure of the Fusarium trypsin-like protease;

Fig. 2 shows the three-dimensional structure of the mammalian trypsin-like proteases isolated from Trypsin (Rattus rattus) (1trm);

Fig. 3 shows the three-dimensional structure of Trypsin (Bo taurus) (2ptn);

Fig. 4 shows the three-dimensional structure of Tonin (Rattus rattus) (1ton);

Fig. 5 shows the three-dimensional structure of Kallikrien A (Sus scrofa) (2pka);

Fig. 6 shows the three-dimensional structure of γ-chymotrypsin (Bos taurus) (2gch);

Fig. 7 shows the three-dimensional structure of Pancreatic Elastase (Sus scrofa) (3est);

Fig. 8 shows the three-dimensional structure of Neutrophil Elastase (Homo sapiens) (1hne);

Fig. 9 shows the three-dimensional structure of Mast Cell Proteinanse (Rattus rattus) (3rp2); and

Fig. 10 shows the three-dimensional structure of the bacterial trypsin-like protease isolated from Trypsin (Streptomyces griseus) (1sgt).

DETAILED DESCRIPTION OF THE INVENTION

In the present description and claims, the following abbreviations are used:

Amino Acids

A	=	Ala	=	Alanine
V	=	Val	=	Valine
L	=	Leu	=	Leucine
I	=	Ile	=	Isoleucine
P	=	Pro	=	Proline
F	=	Phe	=	Phenylalanine
W	=	Trp	=	Tryptophan
M	=	Met	=	Methionine
G	=	Gly	=	Glycine
S	=	Ser	=	Serine
T	=	Thr	=	Threonine
C	=	Cys	=	Cysteine
Y	=	Tyr	=	Tyrosine
N	=	Asn	=	Asparagine
Q	=	Gln	=	Glutamine
D	=	Asp	=	Aspartic Acid
E	=	Glu	=	Glutamic Acid
K	=	Lys	=	Lysine
R	=	Arg	=	Arginine
H	=	His	=	Histidine

Sequence Numbering

In order to simplify the ensuing discussion a sequence numbering based on a structural alignment of the parent trypsin-like Fusarium protease with that of eight mammalian trypsins and one bacterial trypsin have been used. The relationship between the structural sequence numbering used and that of the amino acid sequence shown in SEQ ID No. 2 is apparent from Table 1, below. In Table 1 the residue numbering of γ-chymotrypsin has been used as a reference. A listing of

the structural alignment of the ten trypsin like proteases is shown in Table 2, identified by PDB Codes.

Table 1

Reference-numbering	Amino acid of SEQ ID No. 2	Sequential numbering of SEQ ID No. 2
16	I	1
17	V	2
18	G	3
19	G	4
20	T	5
21	S	6
22	A	7
23	S	8
24	A	9
25	G	10
26	D	11
27	F	12
28	P	13
29	F	14
30	I	15
31	V	16
32	S	17
33	I	18
34	S	19
35	R	20
36	N	21
38	G	22
39	G	23
40	P	24
41	W	25
42	C	26
43	G	27
44	G	28
45	S	29
46	L	30
47	L	31
48	N	32
49	A	33
50	N	34
51	T	35
52	V	36
53	L	37
54	T	38
55	A	39
56	A	40
57	H	41
58	C	42
59	V	43
59a	S	44

59b	G	45
59c	Y	46
60	A	47
61	Q	48
62	S	49
63	G	50
64	F	51
65	Q	52
66	I	53
67	R	54
68	A	55
69	G	56
70	S	57
71	L	58
72	S	59
73	R	60
74	T	61
78	S	62
79	G	63
80	G	64
81	I	65
82	T	66
83	S	67
84	S	68
85	L	69
86	S	70
87	S	71
88	V	72
89	R	73
90	V	74
91	H	75
92	P	76
93	S	77
94	Y	78
95	S	79
98	G	80
99	N	81
100	N	82
101	N	83
102	D	84
103	L	85
104	A	86
105	I	87
106	L	88
107	K	89
108	L	90
109	S	91
110	T	92
111	S	93
112	I	94
113	P	95
114	S	96
115	G	97
116	G	98

117	N	99
118	I	100
119	G	101
120	Y	102
121	A	103
122	R	104
123	L	105
124	A	106
125	A	107
126	S	108
127	G	109
128	S	110
129	D	111
130	P	112
131	V	113
132	A	114
133	G	115
134	S	116
135	S	117
136	A	118
137	T	119
138	V	120
139	A	121
140	G	122
141	W	123
142	G	124
143	A	125
144	T	126
145	S	127
146	E	128
147	G	129
148	G	130
149	S	131
150	S	132
151	T	133
152	P	134
153	V	135
154	N	136
155	L	137
156	L	138
157	K	139
158	V	140
159	T	141
160	V	142
161	P	143
162	I	144
163	V	145
164	S	146
165	R	147
166	A	148
167	T	149
168	C	150
169	R	151
170	A	152

171	Q	153
172	Y	154
173	G	155
174	T	156
175	S	157
176	A	158
177	I	159
177a	T	160
178	N	161
179	Q	162
180	M	163
181	F	164
182	C	165
183	A	166
184	G	167
185	V	168
185b	S	169
185c	S	170
186	G	171
187	G	172
188	K	173
189	D	174
190	S	175
191	C	176
192	Q	177
193	G	178
194	D	179
195	S	180
196	G	181
197	G	182
198	P	183
199	I	184
200	V	185
201	D	186
202	S	187
203	S	188
207	N	189
208	T	190
209	L	191
210	I	192
211	G	193
212	A	194
213	V	195
214	S	196
215	W	197
216	G	198
217	N	199
219	G	200
220	C	201
221	A	202
222	R	203
223	P	204
223a	N	205
224	Y	206

225	S	207
226	G	208
227	V	209
228	Y	210
229	A	211
230	S	212
231	V	213
232	G	214
233	A	215
234	L	216
235	R	217
236	S	218
237	F	219
238	I	220
239	D	221
240	T	222
241	Y	223
242	A	224

Table 2

```
          16  20        30     37abc 40        50       59abc
          |===|=========|======|=====|=========|========|==|
1ton      IVGGYKCEKNSQPWQVAVIN------EYLCGGVLIDPSWVITAAHCY---
2pka      IIGGRECEKNSHPWQVAIYHY----SSFQCGGVLVNPKWVLTAAHCK---
1trm      IVGGYTCQENSVPYQVSLNS-----GYHFCGGSLINDQWVVSAAHCY---
2ptn      IVGGYTCGANTVPYQVSLNS-----GYHFCGGSLINSQWVVSAAHCY---
2gch      IVNGEEAVPGSWPWQVSLQDKT---GFHFCGGSLINENWVVTAAHCG---
3est      VVGGTEAQRNSWPSQISLQYRSGSSWAHTCGGTLIRQNWVMTAAHCV---
1hne      IVGGRRARPHAWPFMVSLQLR----GGHFCGATLIAPNFVMSAAHCV-AN
3rp2      IIGGVESIPHSRPYMAHLDIVTEKGLRVICGGFLISRQFVLTAAHCK---
1sgt      VVGGTRAAQGEFPFMVRLS--------MGCGGALYAQDIVLTAAHCVSGS
sp387     IVGGTSASAGDFPPFIVSISRN----GGPWCGGSLLNANTVLTAAHCVSGY
          |===|=========|======||||==|=========|========|==|
          16  20        30     37abc 40        50       59abc

          62a      70        80        90   95ab  xy  100
          ==||=======|=========|=========|====|||==||====|==
1ton      SN--NYQVLLGRNNLFKDEPFAQRRLVRQSFRHPDYIPL--IPVHDHSND
2pka      ND--NYEVWLGRHNLFENENTAQFFGVTADFPHPGFNLS--ADGKDYSHD
1trm      KS--RIQVRLGEHNINVLEGNEQFVNAAKIIKHPNFD------RKTLNNN
2ptn      KS--GIQVRLGEDNINVVEGNEQFISASKSIVHPSYN------SNTLNND
2gch      VTT-SDVVVAGEFDQGSSSEKIQKLKIAKVFKNSKYN------SLTINND
3est      DRELTFRVVVGEHNLNQNNGTEQYVGVQKIVVHPYWN----TDDVAAGYD
1hne      VNVRAVRVVLGAHNLSRREPTRQVFAVQRIFED-GYD------PVNLLND
3rp2      GR--EITVILGAHDVRKRESTQQKIKVEKQIIHESYN------SVPNLHD
1sgt      GNNTSITATGGVVDLQS--GAAVKVRSTKVLQAPGYN--------GTGKD
sp387     AQS-GFQIRAGSLSRT---SGGITSSLSSVRVHPSYS--------GNNND
          ==||=======|=========|=========|====|||==||====|==
          62a      70        80        90   95ab  xy  100

          103    110       120       130       140    147ab150
```

```
          |======|=========|=========|=========|======|||==|
  1ton    LMLLHLSEPADITGGVKVIDLPT--KEPKVGSTCLASGWGSTNPS--EMV
  2pka    LMLLRLQSPAKITDAVKVLELPT--QEPELGSTCEASGWGSIEPGPDDFE
  1trm    IMLIKLSSPVKLNARVATVALPS--SCAPAGTQCLISGWGNTLSS--GVN
  2ptn    IMLIKLKSAASLNSRVASISLPT--SCASAGTQCLISGWGNTKSS--GTS
  2gch    ITLLKLSTAASFSQTVSAVCLPSASDDFAAGTTCVTTGWGLTRY------
  3est    IALLRLAQSVTLNSYVQLGVLPRAGTILANNSPCYITGWGLTRTNG---Q
  1hne    IVILQLNGSATINANVQVAQLPAQGRRLGNGVQCLAMGWGLLGRNR---G
  3rp2    IMLLKLEKKVELTPAVNVVPLPSPSDFIHPGAMCWAAGWGKTGVR---DP
  1sgt    WALIKLAQPIN----QPTLKIAT--TTAYNQGTFTVAGWGANREG---GS
  sp387   LAILKLSTSIPSGGNIGYARLAASGSDPVAGSSATVAGWGATSEG--GSS
          |======|=========|=========|=========|======|||==|
          103    110       120       130       140    147ab150

          150        0160   169ab     177a180  185abc  190 194
          |========|========|||=======|||=|====||||====|===|
  1ton    VSHDLQCVNIHLLSNEKCI--ETYKDNVT-DVMLCAGEM-EGGKDTCAGD
  2pka    FPDEIQCVQLTLLQNTFCA--DAHPDKVT-ESMLCAGYL-PGGKDTCMGD
  1trm    EPDLLQCLDAPLLPQADCE--ASYPGKIT-DNMVCVGFL-EGGKDSCQGD
  2ptn    YPDVLKCLKAPILSDSSCK--SAYPGQIT-SNMFCAGYL-EGGKDSCQGD
  2gch    TPDRLQQASLPLLSNTNCK--KYWGTKIK-DAMICAGA---SGVSSCMGD
  3est    LAQTLQQAYLPTVDYAICSSSSYWGSTVK-NSMVCAGG--DGVRSGCQGD
  1hne    IASVLQELNVTVV-TSLCR----------RSNVCTLVR-GRQAGVCFGD
  3rp2    TSYTLREVELRIMDEKACV--DYR--YYEYKFQVCVGSP-TTLRAAFMGD
  1sgt    QQRYLLKANVPFVSDAACR--SAYGNRLVANEEICAGYPDTGGVDTCQGD
  sp387   TPVNLLKVTVPIVSRATCR--AQYGTSAITNQMFCAGV-SSGGKDSCQGD
          |========|========|||=======|||=|====||||====|===|
          150        0160   169ab     177a180  185abc  190 194

          195  200 204a   210     218a  223a    230       240
          |====|===||=====|=======||====||======|=========|=
  1ton    SGGPLICD-----GVLQGITSGGAT-PCAKPKTPAIYAKLIKFTSWIKKV
  2pka    SGGPLICN-----GMWQGITSWGHT-PCGSANKPSIYTKLIFYLDWIDDT
  1trm    SGGPVVCN-----GELQGIVSWGY--GCALPDNPGVYTKVCNYVDWIQDT
  2ptn    SGGPVVCS-----GKLQGIVSWGS--GCAQKNKPGVYTKVCNYVSWIKQT
  2gch    SGGPLVCKKN-GAWTLVGIVSWGSS-TCSTS-TPGVYARVTALVNWVQQT
  3est    SGGPLHCLVN-GQYAVHGVTSFVSRLGCNVTRKPTVFTRVSAYISWINNV
  1hne    SGSPLVCN-----GLIHGIASFVRG-GCASGLYPDAFAPVAQFVNWIDSI
  3rp2    SGGPLLCA-----GVAHGIVSYGHP-DAK---PPAIFTRVSTYVPWINAV
  1sgt    SGGPMFRKDNADEWIQVGIVSWGY--GCARPGYPGVYTEVSTFASAIASA
  sp387   SGGPIVDSS----NTLIGAVSWGN--GCARPNYSGVYASVGALRSFIDTY
          |====|===||=====|=======||====||======|=========|=
          195  200 204a   210     218a  223a    230       240
```

```
          242
          |====
  1ton    MKENP
  2pka    ITENP
  1trm    IAAN-
  2ptn    IASN-
  2gch    LAAN-
  3est    IASN-
  1hne    IQ---
  3rp2    IN---
  1sgt    ARTL-
  sp387   A----
          |====
          242
```

In describing trypsin-like protease variants according to the invention, the following nomenclature is used for ease of reference:

Original amino acid(s):position(s):substituted amino acid(s);

the position being indicated in accordance with the structural amino acid numbering apparent from Table 1.

According to this nomenclature, for instance the substitution of glutamic acid for glycine in position 185b is shown as:

Gly 185b Glu or G185bE

a deletion of glycine in the same position is shown as:

Gly 185b * or G185b*

and insertion of an additional amino acid residue such as lysine is shown as:

Gly 185b GlyLys or G185bGK

Where a specific trypsin-like protease contains a "deletion" in comparison with other trypsin-like proteases and an insertion is made in such a position this is indicated as:

* 36 Asp or *36D

for insertion of an aspartic acid in position 36

Multiple mutations are separated by pluses, i.e.:

Arg 170 Tyr + Gly 195 Glu or R170Y+G195E

representing mutations in positions 170 and 195 substituting tyrosine and glutamic acid for arginine and glycine, respectively.

When one or more alternative amino acid residues may be inserted in a given position it is indicated as

G185bE,D,S or

G185bE or G185bD or G185bS

Description of the trypsin-like F. oxysporum protease

The following description of variants of the trypsin-like Fusarium protease of the invention is based on the parent trypsin-like protease derived from the strain of F. oxysporum deposited with the Deutsche Sammlung von Mikroorganismen with the deposit number DSM 2672. It will be understood that functional analogues of said trypsin-like protease as defined above, e.g. other parent trypsin-like Fusarium proteases, may be modified in a similar manner to that described for the trypsin-like F. oxysporum protease, e.g. by modifying similar positions (according to a structural alignment).

The parent trypsin-like F. oxysporum protease is encoded by the DNA sequence shown in the appended Sequence Listing ID No. 1, and the corresponding cDNA sequence and amino acid sequence are shown in SEQ ID Nos. 2 and 3, respectively. The trypsin-like F. oxysporum protease is expressed in the form of an inactive proenzyme also comprising a signal peptide. The mature, active enzyme consists of 224 amino acid residues and has a molecular weight of 22190. The signal peptide is contemplated to be amino acid residues -24 to -8 (according to the rules of von Heijne (1986) and the pro-peptide amino acid residues -7 to -1.

The trypsin-like F. oxysporum protease shows a reversed Arg/Lys specificity compared to that of bovine trypsin, which means that the trypsin-like F. oxysporum protease is more Arg-active than Lys-active.

The three-dimensional structure of the Fusarium trypsin-like protease is shown in Fig. 1 and that of the

mammalian trypsin-like proteases isolated from Trypsin (Rattus rattus) (1trm), Trypsin (Bo taurus) (2ptn), Tonin (Rattus rattus) (1ton), Kallikrien A (Sus scrofa) (2pka), γ-chymotrypsin (Bos taurus) (2gch), Pancreatic Elastase (Sus scrofa) (3est), Neutrophil Elastase (Homo sapiens) (1hne) and Mast Cell Proteinanse (Rattus rattus) (3rp2) in Figs. 2-9, respectively and that of the bacterial trypsin-like protease isolated from Trypsin (Streptomyces griseus) (1sgt) in Fig. 10.

As mentioned above considerable differences in the loop structures of these trypsin-like proteases have been observed. The loop structures of the F. oxysporum trypsin-like protease have been identified and are listed in Table 3. Corresponding loop structures may be identified in other trypsin-like proteases, including the ones illustrated in Figs. 2-10. The amino acid segments listed in this table refer to the positions in the amino acid sequence of the trypsin-like Fusarium protease listed in SEQ ID No. 2. The loops identified in the table may have different backbone conformations.

Table 3

Loop	Segments	Amino acid seq.
I	23-27	SAGDF
II	34-40	RNGG
III	47-50	LNAN
IV	59-64	VSGYAQSGF
V	70-80	SLSRTSGG
VI	91-99	HPSYSGN
VII	110-118	TSIPSGGNI
VIII	125-133	ASGSDPVA
IX	145-151	SEGGSST
X	163-179	SRATCRAQYGTSAITNQ
XI	185-187	GVSSGG
XII	201-208	DSSNT
XIII	215-225	WGNGCARPNYS

The intervening segments which are not mentioned in the above listing form the core of the proteins, and each of these segments have similar backbone conformations for all of the trypsin-like proteases listed in Table 2.

Description of variants of the invention

As explained above the variant of the trypsin-like protease of the invention has one or more characteristic properties, some of which will be explained in detail in the following.

Property i) of the variant of the trypsin-like protease, i.e. the immunological cross reactivity, may be assayed using an antibody raised against or reactive with at least one epitope of the trypsin-like protease comprising the amino acid sequence shown in SEQ ID No. 2. The antibody, which may either be monoclonal or polyclonal, may be produced by methods known in the art, e.g. as described by Hudson et al., 1989. The immunological cross-reactivity may be determined using assays known in the art, examples of which are Western Blotting or radial immunodiffusion assay, e.g. as described by Hudson et al., 1989.

Property ii) of the variant of the trypsin-like protease of the invention, i.e. the homology between the amino acid sequence of the variant and the amino acid sequence shown

in SEQ ID No. 2 is intended to indicate the degree of identity between the two sequences indicating a derivation of the first sequence from the second. In particular, a polypeptide is considered to be homologous to the trypsin-like protease if a comparison of the respective amino acid sequences reveals an identity of greater than about 60%, such as above 70%, 80% or 85%. It will be understood that variants in which a high number of amino acid residues have been modified have a lower degree of identity, whereas variants comprising only few amino acid modifications will show identities of more than 90%, such as more than 95% or even 99% identity. Sequence comparisons can be performed via known algorithms, such as the one described by Lipman and Pearson (1985).

The oligonucleotide probe used in the characterization of the variants of the invention in accordance with property iii) defined above, may suitably be prepared on the basis of the full or partial nucleotide or amino acid sequence shown in SEQ ID No. 1 and 2, encoding or constituting, respectively, the trypsin-like F. oxysporum protease described herein. The hybridization may be carried out under any suitable conditions allowing the DNA sequences to hybridize. For instance, such conditions are hybridization under specified conditions, e.g. involving presoaking in 5xSSC and prehybridizing for 1h at ~40°C in a solution of 20% formamide, 5xDenhardt's solution, 50mM sodium phosphate, pH 6.8, and 50µg of denatured sonicated calf thymus DNA, followed by hybridization in the same solution supplemented with 100µM ATP for 18h at ~40°C, or other methods described by e.g. Sambrook et al., 1989.

In the following specific classes of trypsin-like Fusarium protease variants of the invention having improved properties are described as well as the concepts used for the design of such variants.

Proline stabilization

In order to improve the stability of a trypsin-like Fusarium protease, such as the thermal, storage and/or protease

stability thereof, it may be advantageous to introduce one or more proline residues therein. The improvement obtained by such modification may be explained on the basis of the structure of proteins. The modification of subtilisins by introduction of proline residues therein are described in WO 92/19729, incorporated by reference herein.

Proteases are globular proteins and quite compact due to the considerable amount of folding of the long polypeptide chain. The polypeptide chain essentially consists of the "backbone" and its "side-groups". As the peptide bond is planar, only rotations around the C_α-N axis and the C_α-C' axis are permitted. Rotation around the C_α-N bond of the peptide backbone is denoted by the torsion angle ϕ (phi), rotation around the C_α-C' bond by ψ (psi) [vide e.g. Creighton, T.E. (1984); Proteins; W.H. Freeman and Company, New York]. The choice of the values of these angles of rotation is made by assigning the maximum value of +180° (which is identical to -180°) to the maximally extended chain. In the fully extended polypeptide chain, the N, C_α and C' atoms are all "trans" to each other. In the "cis" configuration, the angles ϕ and ψ are assigned the value of 0°. Rotation from this position around the bonds so that the atoms viewed behind the rotated bond move "counterclockwise" is assigned negative values by definition, those "clockwise" are assigned positive values. Thus, the values of the torsion angles lie within the range -180° to +180°.

Since the C_α-atoms are the swivel point for the chain, the side-groups (R-groups) associated with the C_α-atoms become extremely important with respect to the conformation of the molecule. The conformation, which defines the participation of the secondary and tertiary structures of the polypeptide chains in moulding the overall structure of the protein, is of prime importance to the specific structure of the protein and contributes greatly to the unique catalytic properties (i.e. activity and specificity) of the enzyme and its stability.

Proline residues have a reduced degree of rotational freedom around the N-C_α bond compared to other types of amino

acids, because the proline sidechain connects back to the amide nitrogen. This connectivity usually restricts the ϕ-angles of proline residues to a narrow interval around -60°.

The equilibrium between the unfolded and folded state of a protein is to a large extent governed by the entropy difference between the two states, and consequently it is envisaged that the trypsin-like Fusarium protease can be stabilized by reducing the number of different conformations that are accessible in the unfolded state. Introduction of proline residues for other residues in the protein sequence generally reduces the entropy of the unfolded state, due to the restricted rotational freedom for proline residues. For such substitutions to have effect on the stability of the protein, they must be compatible with the structure of the protein in the folded state, that is, the substituted residues must have ϕ-angles in the folded state that are in the allowed interval for prolines, and the introduced prolines must not cause an energetically unfavourable packing of the protein atoms.

In accordance with the above theoretical explanation trypsin-like Fusarium protease variants are contemplated, in which a naturally occurring amino acid residue (other than proline) of the amino acid sequence of the parent trypsin-like protease has been substituted with a proline residue at one or more positions, at which positions(s) the dihedral angles ϕ (phi) and ψ (psi) constitute values within the intervals [$-90° < \phi < -40°$ and $-180° < \psi < 180°$], preferably within the intervals [$-90° < \phi < -40°$ and $120° < \psi < 180°$] or [$-90° < \phi < -40°$ and $-50° < \psi < 10°$], and which position(s) is/are not located in regions in which the protease is characterized by possessing α-helical or ß-sheet structure.

The stabilized trypsin-like Fusarium protease variants according to this embodiment of the invention may be prepared by subjecting the trypsin-like Fusarium protease to analysis for secondary structure, identifying residues in the protease having dihedral angles ϕ (phi) and ψ (psi) confined to the intervals [$-90° < \phi < -40°$ and $-180° < \psi < 180°$], preferably the inter-

vals [-90°<ϕ<-40° and 120°<ψ<180°] or [-90°<ϕ<-40° and -50°<ψ<10°], excluding residues located in regions in which the trypsin-like Fusarium protease is characterized by possessing α-helical or ß-sheet structure, if a proline residue is not already at the identified position(s), substitution of the naturally occurring amino acid residue with a proline residue at the identified position(s), preferably by site directed mutagenesis of a gene encoding the trypsin-like Fusarium protease, and subsequent expression and recovery of the resulting trypsin-like Fusarium protease variant as described in detail below.

The analysis of the parent trypsin-like Fusarium protease for secondary structure may be performed on the basis of the atomic structure of the protease, which, e.g. may be determined by X-ray diffraction techniques. X-ray diffraction techniques are described by e.g. Hendrickson, W.A [X-ray diffraction; in Protein Engineering (Ed: Oxender, D.L and Fox, C.F.), ch. 1; Alan R. Liss, Inc. (1987)] and Creighton, T.E., supra, ch. 6.

When the atomic structure has been determined, it is possible to compute dihedral angles from the atomic coordinates. Moreover, it is possible to assign secondary structure elements. The secondary structure elements are defined on the basis of hydrogen bindings. Cooperative secondary structure is recognized as repeats of the elementary hydrogen-bonding patterns "turn" and "bridge". Repeating turns are "helices", repeating bridges are "ladders", connected ladders are "sheets".

Analysis for secondary structure elements requires a computerized compilation of structure assignments and geometrical features extracted from atomic coordinates. The conventional method to elucidate the secondary structure of a protein, based on its atomic coordinates, is described by Kabsch, W. and Sander, C. [Biopolymers (1983) **22** 2577-2637]. In this article an algorithm for extracting structural features from

the atomic coordinates by a pattern-recognition process is provided. First, H-bonds are identified based on electrostatic interactions between pairs of H-bonding groups. Next, the patterns of H-bonding are used to define secondary structure elements such as turns (T), bends (S), bridges (B), helices (G,H,I), ß-ladders (E) and ß-sheets (E).

A computer program DSSP (Define Secondary Structure of Proteins), enabling the computation of Kabsch & Sander files and written in standard PASCAL, is available from the Protein Data Bank, Chemistry Dept., Brookhaven National Laboratory, Upton, N.Y. 11973.

After the dihedral angles ϕ (phi) and ψ (psi) for the amino acids have been calculated, based on the atomic structure in the crystalline proteases, it is possible to select position(s) which has/have dihedral phi and psi angles favourable for substitution with a proline residue. The aliphatic side chain of proline residues is bonded covalently to the nitrogen atom of the peptide group. The resulting cyclic five-membered ring consequently imposes a rigid constraint on the rotation about the N-C_α bond of the peptide backbone and simultaneously prevents the formation of hydrogen bonding to the backbone N-atom. For these structural reasons, prolines are generally not compatible with α-helical and ß-sheet secondary conformations. Due to the same rotational constraint about the C_α-N bond, and due to the requirement that neighbouring amino acids in the chain are not perturbed, the magnitudes of the dihedral angles phi and psi (and in particular phi) are confined to limited intervals for proline residues in polypeptides. The dihedral angles for proline residues in polypeptides are almost exclusively within the intervals [$-90° < \phi < -40°$ and $-180° < \psi < 180°$], preferably the intervals [$-90° < \phi < -40°$ and $120° < \psi < 180°$] or [$-90° < \phi < -40°$ and $-50° < \psi < 10°$]. In this context, both cis- and trans-proline residues are considered.

For the reasons stated above it is preferred that the amino acid residue(s) to be substituted with proline is a hydrophilic or a small hydrophobic amino acid residue, in

particular one selected from the amino acid residues A, D, E, K, G, Q, R, S, T, N and V.

Studies of the three-dimensional structure of the trypsin-like F. oxysporum protease have revealed the following positions occupied by amino acid residues, of which one or more may advantagously be substituted with a proline residue: 21, 24, 30, 46, 49, 71, 90, 111, 114, 121, 124, 125, 126, 132, 135, 146, 148, 174, 175, 176, 178, 185b, 185c, 202, 208, and 209 (using the amino acid numbering defined in Table 1, above). It is believed that substitution for proline in one or more of these positions may result in trypsin-like Fusarium protease variants with an improved stability.

Specific variants of the trypsin-like F. oxysporum protease comprise one or more of the following substitutions: A24P, A49P, V90P, S111P, A124P, A125P, S126P, A132P, S135P, T174P, S175P, S185bP, S185cP, S202P.

Stabilization by introduction of a disulphide-bridge

Stabilization of a given protein, especially as concerns thermostabilization, may be achieved by covalently binding two regions in the protein that are far apart in sequence but close in space. Such binding may be performed by the introduction of a disulphide-bridge in the protein, i.e. by introducing one or more cysteine residues capable of binding to each other or to other cysteine residues present in the protein.

In accordance with this embodiment, the invention relates to a variant of a trypsin-like Fusarium protease, in which an amino acid residue different from cysteine of the amino acid sequence of the parent trypsin-like protease has been substituted with a cysteine residue in such a manner that the introduced cysteine residue together with another cysteine residue present in the parent protease or introduced therein form a disulphide bridge.

Positions in which SS-bridges may be introduced may be identified by comparing the structure of the trypsin-like

Fusarium protease with the structures of the homologous trypsins listed in Table 2. By such comparison the following residue pair positions have been identified, between which a disulphide bridge can be introduced:

22+157, 129+232, 136+201.

Specific variants of the trypsin-like F. oxysporum protease include:

(i) A22C+D26S+K157C,

(ii) A136C+D201C, and

(iii) A124P+A125T+S126*+G127*+D129C+P130A+G232C

In (i) the substitution D26S is performed to allow for the introduction of Cys in position 157 due to a salt bridge between D26 and K157.

In (iii) due account has been taken of positions 124 to 128 by shortening loop no. VII.

Stabilization by modification of Asn-Gly Pairs

It is known that at alkaline pH, the sidechain of Asn may interact with the NH group of a sequential neighbouring amino acid to form an isoAsp residue where the backbone goes through the Asp sidechain. This will leave the backbone more vulnerable to proteolysis. The deamidation is much more likely to occur if the residue that follows is a Gly. Changing the Asn in front of the Gly or the Gly will prevent this from happening and thus improve the stability, especially as concerns thermo- and storage stability.

The invention consequently further relates to a trypsin-like Fusarium protease variant, in which either or both residues of any of the Asn-Gly sequence appearing in the amino acid sequence of the parent trypsin-like protease is/are deleted or substituted with a residue of a different amino acid.

The Asn and/or Gly residue may, for instance, be substituted with a residue of an amino acid selected from the group consisting of A, Q, S, P, T and Y.

More specifically, any of the Asn or Gly residues of the Asn-Gly occupying positions 36+38 and/or 217+219 of the parent trypsin-like protease (using the amino acid numbering defined in Table 1, above) may be deleted or substituted with a residue of an amino acid selected from the group consisting of A, Q, S, P, T and Y. In this connection it should be noted that there are jumps in the sequence numbering of trypsin-like Fusarium protease, and the two positions mentioned there are therefore sequential.

Specific variants of SP387 are:
N36S, G38S, N217S,Y, G219S

Introduction of protease resistant loops

In a further embodiment the present invention relates to a trypsin-like Fusarium protease variant, in which one or more amino acid residues present in or constituting a loop structure of the parent trypsin-like protease susceptible to cleavage by a proteolytic enzyme is/are deleted or replaced with one or more amino acid residues so as to obtain a modified loop structure having an improved proteolytic stability.

It is of particular interest to modify loop structures of a trypsin-like Fusarium protease which are susceptible to cleavage by a subtilisin or trypsin.

Thus, subtilisin and/or trypsin-resistant loops may be identified in other proteins and substituted for loop(s) of a parent trypsin-like Fusarium protease having a lower proteolytic stability. The substitution may be performed either by substituting the entire loop(s) or by substitution one or more amino acid residues of a loop of the parent trypsin-like Fusarium protease so as to obtain a modified loop having a higher protease stability than the original loop. The amino

acid residues to be substituted may be identified by comparing the amino acid sequence of the loop of the parent trypsin-like Fusarium protease with that of the "foreign" more protease-resistant loop.

More specifically, trypsin-like Fusarium protease variants according to this embodiment comprises variants in which one or more amino acid substitutions is/are made so that the modified loop structure comprises at least one of the amino acid segments GAAG, GARG, YPGS, YPRS, HNRG, YTGN, ISSE, NNAG, SFIN, DQNG, ASFS, SRGV, LDTG, YYAA, INDI, WYFG, SIEN, GSTY, DSTN, PDLR, LDTG, GNRY, SGVM, RYPS, NGLV, SFSI, LGSP, RASF, VPWG, PDLN, SFVP, PDYR, PRLP, TVLP, IGTC, TGGT, TNKL, VGDV, IGVL, GSTY, RYAN, PNIP, or TLVP are contemplated to have an improved proteolytic stability.

Other specific variants according to this embodiment of the invention include a variant, in which loop II of the trypsin-like protease comprising the peptide sequence SRNGGP is substituted with loop II of the trypsin 2ptn isolated from Bos taurus comprising the peptide sequence NSGYH, as follows:

S34N+R35S+N36*+G39Y+P40H, and/or

loop IV of the trypsin-like protease comprising the peptide sequence VSGYAQSGF is substituted with loop IV of the trypsin 2ptn isolated from Bos taurus comprising the peptide sequence YKSGI, as follows:

V59Y+S59a*+G59b*+Y59c*+A60K+Q61S+S62*+F64I, and/or

loop IV of the trypsin-like protease comprising the peptide sequence VSGYAQSGF is substituted with loop IV of the trypsin 1trm isolated from Rattus rattus comprising the peptide sequence YKSRI, as follows:

V59Y+S59a*+G59b*+Y59c*+A60K+Q61S+S62*+G63R+F64I, and/or

loop VI of the trypsin-like protease comprising the peptide sequence HPSYSGN is substituted with loop VI of the trypsin 1trm isolated from Rattus rattus comprising the peptide sequence HPNFDRKTL, as follows:

S93N+F94Y+S95D+*96R+*97K+G98T+N99L.

Removal of autoproteolysis sites

According to a further aspect of the invention autoproteolysis sites may be removed by changing the amino acids at the autoproteolysis site. Since the trypsin-like F. oxysporum protease cleaves at Lys and Arg residues it is preferred to modify such residues of a parent trypsin-like Fusarium protease having the same or a similar specificity, preferably by substituting with a non or less positively charged amino acid residue. The non or less positively charged amino acid residue may be selected from the group consisting of K, S, V, P, E, D, N, Q, A and G; the amino acid residues K, S, V, or P being particularly preferred.

Since the parent trypsin-like F. oxysporum protease is specific mostly towards Arg and to a minor extend towards Lys residues, the modification of this parent trypsin-like protease may preferably be made by changing Arg to another amino acid residue (including Lys) or by changing Arg or Lys to a non or less positively charged amino acid as defined above.

The following residue positions of the trypsin-like F. oxysporum protease have been found to contain Lys or Arg residues:

35, 67, 73, 89, 107, 122, 157, 165, 169, 188, 222, 235, using the structural sequence numbering defined in Table 1.

Of these potential autoproteolysis sites, residues 35, 122 and 169 have been found to be primary autoproteolysis sites, and positions 73 and 89 have been found to be secondary autoproteolysis sites. Examples of specific modifications are:

R35K,S
R122K,V,P
R169K,V
R122K,V,P+A124P
R122V,P+A124P+T208V.

Alternatively autoproteolysis can be prevented by changing the amino acid residue occupying the position following the Arg or Lys residue in question to Pro. For instance, this may be done in the positions 90 and/or 123 (according to the structural amino acid numbering defined in Table 1), as follows:

V90P, L123P.

Removal of critical oxidation sites

In order to increase the stability of the trypsin-like Fusarium protease it may be advantageous to substitute critical oxidation sites, such as methionines, with other amino acid residues which are not subject to oxidation.

Accordingly, in a further embodiment the present invention relates to a trypsin-like Fusarium protease variant, in which one or more amino acid residues susceptible to oxidation is/are replaced with another amino acid residue less susceptible to oxidation. The amino acid residue less susceptible to oxidation may for instance be selected from the group consisting of A, E, N, Q, I, L, S and K.

Specific variants of the trypsin-like F. oxysporum protease comprises one of the following substitutions:

M180N,Q,E,K, using the sequence numbering of Table 1.

Modification of tryptophan residues

In order to stabilize the protein it may be advantageous to replace or delete tryptophan residues at the surface of the protein, e.g. as described in US 5,118,623. The tryptophan

residues may advantageously be substituted for F, T, Q or G. Thus, in a further embodiment the invention relates to a trypsin-like F. oxysporum variant comprising one or more of the following substitutions:

W41F,T,Q,G
W215F,T,Q,G

Introduction of Glycosylation Sites

The concept is to introduce N-glycosylation sites in loops subject to proteolysis in general and autoproteolysis in particular.

More specifically, in accordance with this embodiment the invention relates to a trypsin-like Fusarium protease variant, in which a N-glycosylation site has been introduced at an amino acid residue X located in a loop structure on the surface of the protein subject to proteolysis by changing the sequence segment X-Y-Z to Asn-Y1-Ser or Asn-Y1-Thr, provided that Y1 is different from Gly, so as to confer to the variant an improved proteolysis resistance.

The introduction of a N-glycosylation site at residue X will make the segment target for proteins that N-glycosylate the amide nitrogen of the Asn residue introduced. The residues X, Y, Y1, Z can be any residue, except Y1 should not be Gly, because this would create an Asn-Gly sequence the introduction of which may result in a less stable variant. For Asn in the changed sequence segment to become glycosylated, the changed sequence segment must lie on the protein surface such that it can be recognized by glycosylation proteins.

In the parent trypsin-like F. oxysporum protease, a N-glycosylation site can be introduced at the following positions:

17, 18, 20, 21, 23, 24, 25, 34, 35, 36, 38, 39, 48, 49, 50, 59a, 59b, 59c, 60, 61, 62, 63, 69, 70, 71, 72, 73, 74, 78, 79, 80, 81, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 98, 99, 109, 110, 111, 113, 114, 115, 116, 117, 118, 121, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 143, 144, 145,

146, 147, 148, 149, 150, 151, 152, 153, 154, 163, 164, 165, 173, 174, 175, 176, 177a, 178, 184, 185, 185b, 187, 188, 201, 202, 203, 207, 215, 216, 217, 221, 222, 223, 224, 225, 233, 234, 235 and/or 236. The numbers are given in accordance with the amino acid numbering defined in Table 1 and refer to the starting position of the target segment. Thus, for instance, position "17" indicates the segment 17-18-19, "18" the segment 18-19-20 etc.

The technique may specifically be applied to primary autoproteolysis sites containing Lys or Arg residues, that is cases where Y = Lys or Arg. In cases where the Lys or Arg is considered important for the functioning of the protein, for instance where a reduction in pI is undesirable, the auto--proteolysis sites cannot be removed by substituting Lys and Arg to other residues where the parent trypsin-like Fusarium protease has low specificity. By introducing N-glycosylation sites on residues that are sequential neighbours to exposed Lys and Arg residues one may reduce the binding affinity of parent trypsin-like Fusarium protease for the auto-proteolysis site, and thereby prevent auto-proteolysis, while maintaining the Lys or Arg residues. That is Y=Y1=Lys or Arg.

In this connection, introduction of a N-glycosylation site by substitution of an amino acid residue of the parent trypsin-like F. oxysporum protease occupying any of the positions 34, 36, 72, 74, 88, 90, 121, 123, 164, 170, 187 (using the sequence numbering defined in Table 1) may be advantageous.

Examples of trypsin-like F. oxysporum variants comprises one or more of the following substitutions:

S34N+R35R,K+N36S,T
N36N+G38S+G39S,T
S72N+R73R,K+T74T,S
V88N+R89R,K+V90S,T
A121N+R122R,K+L123S,T
L123N+A224A+A225S,T
S164N+R165R,K+A166S,T

Variants with improved wash performance

The ability of an enzyme to catalyse the degradation of various naturally occurring substrates present on the objects to be cleaned during e.g. wash is often referred to as its washing ability, washability, detergency, or wash performance. The present invention devices trypsin-like Fusarium protease variants having an improved wash performance as compared to that of the parent trypsin-like protease.

Variants with improved stability against inactivation by peroxidase or pHBS

It is possible that many, if not all tyrosines are chemically modifed by peroxidases and/or pHBS (p-hydroxy benzene sulfonate). This kind of modification may cause a dramatic decrease in the isoelectric point of trypsin-like Fusarium proteases. To prevent such modification Tyr residues may be modified to other amino acids, preferably a hydrophobic amino acid when the Tyr is burried in the interior of the protein or a hydrophilic amino acid when the Tyr is exposed on the protein surface.

In the parent trypsin-like F. oxysporum protease, the following tyrosine residues may be modified:
59c, 94, 120, 172, 224, 228, 241.

Examples of specific trypsin-like F. oxysporum variants comprises on or more of the following substitutions:
Y59cN,Q,S,A,F
Y94F,W
Y120T,F,S,A,V
Y172F,W,H
Y224T,S,K,N,D,F
Y228F
Y241T,S,V,F.

Variants with raised/lowered pI

The concept is to alter the pI for the protein such that it approaches the pH of the detergent formulation. The pI can be raised by changing negatively charged or neutral amino acids to positively charged amino acids or by changing positively charged residues to more positively charged residues. The pI can be lowered by changing positively charged or neural amino acids to negatively charged amino acids or by changing negatively charged amino acids to more negatively charged amino acids.

Accordingly, in accordance with this embodiment the invention relates to a trypsin-like Fusarium protease variant, in which the net electrostatic charge of the parent trypsin-like protease has been changed by deleting or substituting one or more negatively charged amino acid residues by neutral or positively charged amino acid residue(s), and/or by substituting one or more neutral amino acid residues by positively or negatively charged amino acid residue(s), and/or by deleting or substituting one or more positively charged amino acid residue(s) by neutral or negatively charged amino acid residue(s), thereby obtaining variant which either has a lower or higher pI as compared to the pI of its parent protease.

In order to have any effect on the pI, the positions suited for substitution should be located on the protein surface. It is preferred that the amino acid substitutions result in a variant protease having a pI just below the pH of the detergent.

In particular, an amino acid residue located in one or more positions of the parent trypsin-like Fusarium protease and exposed at the surface of the molecule may be substituted:
17, 18, 20, 21, 23, 24, 25, 26, 27, 34, 35, 36, 38, 39, 40, 41, 48, 49, 50, 59a, 59b, 59c, 60, 61, 62, 63, 65, 67, 69, 70, 71, 72, 73, 74, 75, 76, 78, 79, 80, 81, 82, 83, 86, 87, 88, 89, 90, 91, 93, 94, 95, 98, 99, 100, 101, 109, 110, 111, 113, 114, 115, 116, 117, 118, 119, 120, 122, 123, 124, 125, 126, 127, 128,

129, 130, 131, 132, 133, 134, 135, 137, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 156, 157, 159, 161, 163, 164, 165, 166, 167, 169, 170, 171, 173, 174, 175, 176, 177a, 178, 179, 184, 185, 185b, 185c, 187, 188, 189, 192, 201, 202, 207, 208, 210, 215, 216, 217, 219, 221, 222, 223, 223a, 224, 225, 226, 233, 234, 235, 236, 237, 239, 240, 241 and/or 242, the numbering being in accordance with the structural amino acid numbering defined in Table 1.

Altered polarity in active site

It has been shown for rat trypsin that the catalytic activity towards the substrate Succinly-Ala-Phe-Lys-AMC can be increased with a factor of 1.84 (K^{mut}_{cat}/K_{cat}) by changing the active site polarity around Asp102 through a substitution of the conserved Ser214 to Ala (McGrath et al., Biochemistry (1992) 31, 3059-3064). The substitution S214A causes the Ser214-OH group to be replaced by a water molecule, which again causes the change in polarity around Asp102.

In accordance herewith, the invention further relates to a trypsin-like Fusarium protease, in which one or more amino acids around the active site has been substituted with an other amino acid sequence so as to obtain a change in the polarity around the active site. In this context it is contemplated that substituting the serine at position 214 in a parent trypsin-like Fusarium protease with another amino acid may make the enzyme more active.

An example of a variant of the trypsin-like F. oxysporum protease according to this embodiment of the invention comprises the substitution S214A.

Removal of Glycosylation Sites

The concept is to remove critical or potential N-glycosylation sites near the binding cleft region as they might interfere with binding of substrate. For instance, a variant of a trypsin-like F. oxysporum protease in which the N-glycosylation site 223a has been removed may result in an

improved substrate binding. More specifically the following substitutions are contemplated: N223aS,G,R,K

Surface Loops near the Active Site

It has been shown for trypsin (Hedstrom, L et al.: Science (1992) **255**, pp. 1249-1253) that surface loops near the active site influence its specificity and catalytic rate. Thus in a further aspect the present invention relates to a trypsin-like protease variant improved by substituting any of its surface loops near the active site with the corresponding surface loop from the trypsin-like F. oxysporum protease disclosed herein.

In particular substituting any of the active site surface loops II, IV, VI, IX, X (in particular pos. 171-178 thereof), and XIII (in particular pos. 215-221 thereof) defined above in any trypsin-like protease with a loop that is a) at least 60% homologous to the corresponding loop in the trypsin-like F. oxysporum protease, or b) reacts with antibodies raised against the corresponding loop of the trypsin-like F. oxysporum protease and which recognizes the trypsin-like F. oxysporum protease, is believed to improve properties of the trypsin-like protease in question.

Of course the loop to be inserted may show a higher homology to that of the trypsin-like F. oxysporum protease, for instance a homology of at least 80%, such as at least 85%, 90% or even at least 95% with that of the corresponding loop structure of the trypsin-like F. oxysporum protease.

The loop structure of a trypsin-like protease which "corresponds to" a given loop structure of the trypsin-like F. oxysporum protease may easily be determined by comparison of the three-dimensional structures of the trypsin-like protease in question with that of the trypsin-like F. oxysporum protease.

The loop structure to be inserted may either be provided by substituting one or more amino acid residues of the

parent loop structure so as to result in the desired modification, or by substituting the entire loop.

The properties a) and b) mentioned above is intended to be understood in a similar manner to that of properties i) and ii) defined above.

It is believed that the parent trypsin-like protease to be modified in accordance with this aspect of the invention may be derived from a variety of sources including mammals, vertebrates, insects, microorganism and the like. Examples of mammalian and bacterial trypsin-like proteases are apparent from Table 2 above.

It should be noted that, according to the invention, any one of the modifications of the amino acid sequence indicated above for the various classes of trypsin-like protease variants may be combined with any one of the other modifications mentioned above where appropriate.

Methods of preparing trypsin-like protease variants

Several methods for introducing mutations into genes are known in the art. After a brief discussion of cloning trypsin-like protease-encoding DNA sequences (which for instance encode functional analogous of the trypsin-like F. oxysporum protease disclosed herein), methods for generating mutations at specific sites within the trypsin-like protease-encoding sequence will be discussed.

Cloning a DNA sequence encoding a trypsin-like protease

The DNA sequence encoding a parent trypsin-like protease may be isolated from any cell or microorganism producing the trypsin-like protease in question by various methods, well known in the art. First a genomic DNA and/or cDNA library should be constructed using chromosomal DNA or messenger RNA from the organism that produces the trypsin-like protease to be studied. Then, if the amino acid sequence of the trypsin-like protease is known, homologous, labelled oligonucleotide probes may be synthesized and used to identify

trypsin-like protease-encoding clones from a genomic library prepared from the organism in question. Alternatively, a labelled oligonucleotide probe containing sequences homologous to a known trypsin-like protease could be used as a probe to identify trypsin-like protease-encoding clones, using hybridization and washing conditions of lower stringency.

Yet another method for identifying trypsin-like protease-producing clones would involve inserting fragments of genomic DNA into an expression vector, such as a plasmid, transforming trypsin-like protease-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing a substrate for the trypsin-like protease thereby allowing clones expressing the trypsin-like protease to be identified.

Alternatively, the DNA sequence encoding the enzyme may be prepared synthetically by established standard methods, e.g. the phosphoamidite method described by S.L. Beaucage and M.H. Caruthers, Tetrahedron Letters 22, 1981, pp. 1859-1869, or the method described by Matthes et al., The EMBO J. 3, 1984, pp. 801-805. According to the phosphoamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

Finally, the DNA sequence may be of mixed genomic and synthetic, mixed synthetic and cDNA or mixed genomic and cDNA origin prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate), the fragments corresponding to various parts of the entire DNA sequence, in accordance with standard techniques. The DNA sequence may also be prepared by polymerase chain reaction (PCR) using specific primers, for instance as described in US 4,683,202 or R.K. Saiki et al., Science 239, 1988, pp. 487-491.

Site-directed mutagenesis

Once a trypsin-like protease-encoding DNA sequence has been isolated, and desirable sites for mutation identified,

mutations may be introduced using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites; mutant nucleotides are inserted during oligonucleotide synthesis. In a specific method, a single-stranded gap of DNA, bridging the trypsin-like protease-encoding sequence, is created in a vector carrying the trypsin-like protease gene. Then the synthetic nucleotide, bearing the desired mutation, is annealed to a homologous portion of the single-stranded DNA. The remaining gap is then filled in with DNA polymerase I (Klenow fragment) and the construct is ligated using T4 ligase. A specific example of this method is described in Morinaga et al., (1984, Biotechnology 2:646-639). U.S. Patent number 4,760,025, by Estell et al., issued July 26, 1988, discloses the introduction of oligonucleotides encoding multiple mutations by performing minor alterations of the cassette, however, an even greater variety of mutations can be introduced at any one time by the Morinaga method, because a multitude of oligonucleotides, of various lengths, can be introduced.

Another method of introducing mutations into trypsin-like protease-encoding sequences is described in Nelson and Long, Analytical Biochemistry 180, 1989, pp. 147-151. It involves the 3-step generation of a PCR fragment containing the desired mutation introduced by using a chemically synthesized DNA strand as one of the primers in the PCR reactions. From the PCR-generated fragment, a DNA fragment carrying the mutation may be isolated by cleavage with restriction endonucleases and reinserted into an expression plasmid.

Expression of trypsin-like protease variants

According to the invention, a mutated trypsin-like protease-coding sequence produced by methods described above, or any alternative methods known in the art, can be expressed, in enzyme form, using an expression vector which typically includes control sequences encoding a promoter, operator,

ribosome binding site, translation initiation signal, and, optionally, a repressor gene or various activator genes.

The recombinant expression vector carrying the DNA sequence encoding a trypsin-like protease variant of the invention encoding may be any vector which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid, a bacteriophage or an extrachromosomal element, minichromosome or an artificial chromosome. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the DNA sequence encoding a trypsin-like protease variant of the invention, especially in a bacterial host, are the promoter of the lac operon of E.coli, the Streptomyces coelicolor agarase gene dagA promoters, the promoters of the Bacillus licheniformis α-amylase gene (amyL), the promoters of the Bacillus stearothermophilus maltogenic amylase gene (amyM), the promoters of the Bacillus Amyloliquefaciens α-amylase (amyQ), the promoters of the Bacillus subtilis xylA and xylB genes etc. For transcription in a fungal host, examples of useful promoters are those derived from the gene encoding A. oryzae TAKA amylase, Rhizomucor miehei aspartic proteinase, A. niger neutral α-amylase, A. niger acid stable α-amylase, A. niger glucoamylase, Rhizomucor miehei lipase, A. oryzae alkaline

protease, A. oryzae triose phosphate isomerase or A. nidulans acetamidase.

The expression vector of the invention may also comprise a suitable transcription terminator and, in eukaryotes, polyadenylation sequences operably connected to the DNA sequence encoding the recombinant protease of the invention. Termination and polyadenylation sequences may suitably be derived from the same sources as the promoter.

The vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the dal genes from B.subtilis or B.licheniformis, or one which confers antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracyclin resistance. Furthermore, the vector may comprise Aspergillus selection markers such as amdS, argB, niaD and sC, a marker giving rise to hygromycin resistance, or the selection may be accomplished by co-transformation, e.g. as described in WO 91/17243.

While intracellular expression may be advantageous in some respects, e.g. when using certain bacteria as host cells, it is generally preferred that the expression is extracellular. As mentioned above the trypsin-like F. oxysporum protease comprising the amino acid sequence shown in the SEQ ID No. 2 comprises a preregion permitting secretion of the expressed protease into the culture medium. If desirable, this preregion may be substituted with a different preregion or signal sequence, convenient accomplished by substitution of the DNA sequences encoding the respective preregions.

The procedures used to ligate the DNA construct of the invention encoding a trypsin-like protease variant, the promoter, terminator and other elements, respectively, and to insert them into suitable vectors containing the information

necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al. (1989)).

The cell of the invention either comprising a DNA construct or an expression vector of the invention as defined above is advantageously used as a host cell in the recombinant production of a trypsin-like protease variant of the invention. The cell may be transformed with the DNA construct of the invention encoding the variant, conveniently by integrating the DNA construct in the host chromosome. This integration is generally considered to be an advantage as the DNA sequence is more likely to be stably maintained in the cell. Integration of the DNA constructs into the host chromosome may be performed according to conventional methods, e.g. by homologous or heterologous recombination. Alternatively, the cell may be transformed with an expression vector as described below in connection with the different types of host cells.

The cell of the invention may be a cell of a higher organism such as a mammal or an insect, but is preferably a microbial cell, e.g. a bacterial or a fungal (including yeast) cell.

Examples of suitable bacteria are grampositive bacteria such as Bacillus subtilis, Bacillus licheniformis, Bacillus lentus, Bacillus brevis, Bacillus stearothermophilus, Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus coagulans, Bacillus circulans, Bacillus lautus, Bacillus megaterium, Bacillus thuringiensis, or Streptomyces lividans or Streptomyces murinus, or gramnegative bacteria such as E.coli. The transformation of the bacteria may for instance be effected by protoplast transformation or by using competent cells in a manner known per se.

The yeast organism may favourably be selected from a species of Saccharomyces or Schizosaccharomyces, e.g. Saccharomyces cerevisiae. The filamentous fungus may advantageously belong to a species of Aspergillus, e.g. Aspergillus oryzae or Aspergillus niger. Fungal cells may be transformed by a process involving protoplast formation and transformation of the

protoplasts followed by regeneration of the cell wall in a manner known per se. A suitable procedure for transformation of Aspergillus host cells is described in EP 238 023.

In a yet further aspect, the present invention relates to a method of producing a trypsin-like protease variant of the invention, which method comprises cultivating a host cell as described above under conditions conducive to the production of the protease and recovering the protease from the cells and/or culture medium.

The medium used to cultivate the cells may be any conventional medium suitable for growing the host cell in question and obtaining expression of the protease variant of the invention. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection).

The trypsin-like protease variant secreted from the host cells may conveniently be recovered from the culture medium by well-known procedures including separating the cells from the medium by centrifugation or filtration, and precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, followed by chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

Detergent Compositions

According to the invention, the protease variant may typically be a component of a detergent composition. As such, it may be included in the detergent composition in the form of a non-dusting granulate, a stabilized liquid, or a protected enzyme. Non-dusting granulates may be produced, e.g., as disclosed in US 4,106,991 and 4,661,452 (both to Novo Industri A/S) and may optionally be coated by methods known in the art. Examples of waxy coating materials are poly(ethylene oxide) products (polyethyleneglycol, PEG) with mean molecular weights of 1000 to 20000; ethoxylated nonylphenols having from 16 to 50 ethylene oxide units; ethoxylated fatty alcohols in which the

alcohol contains from 12 to 20 carbon atoms and in which there are 15 to 80 ethylene oxide units; fatty alcohols; fatty acids; and mono- and di- and triglycerides of fatty acids. Examples of film-forming coating materials suitable for application by fluid bed techniques are given in patent GB 1483591. Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid according to established methods. Other enzyme stabilizers are well known in the art. Protected enzymes may be prepared according to the method disclosed in EP 238,216.

The detergent composition of the invention may be in any convenient form, e.g. as powder, granules, paste or liquid. A liquid detergent may be aqueous, typically containing up to 70% water and 0-30% organic solvent, or nonaqueous.

The detergent composition comprises one or more surfactants, each of which may be anionic, nonionic, cationic, or zwitterionic. The detergent will usually contain 0-50% of anionic surfactant such as linear alkylbenzenesulfonate (LAS), alpha-olefinsulfonate (AOS), alkyl sulfate (fatty alcohol sulfate) (AS), alcohol ethoxysulfate (AEOS or AES), secondary alkanesulfonates (SAS), alpha-sulfo fatty acid methyl esters, alkyl- or alkenylsuccinic acid, or soap. It may also contain 0-40% of nonionic surfactant such as alcohol ethoxylate (AEO or AE), carboxylated alcohol ethoxylates, nonylphenol ethoxylate, alkylpolyglycoside, alkyldimethylamine oxide, ethoxylated fatty acid monoethanolamide, fatty acid monoethanolamide, or polyhydroxy alkyl fatty acid amide (e.g. as described in WO 92/06154).

The detergent composition may additionally comprise one or more other proteases as well as one or more other enzymes conventionally used in detergent compositions, such as an amylase, a lipase, a cutinase, a cellulase, a peroxidase, and/or an oxidase, e.g., a laccase.

The detergent may contain 1-65% of a detergent builder or complexing agent such as zeolite, diphosphate, triphosphate,

phosphonate, citrate, nitrilotriacetic acid (NTA), ethylenediaminetetraacetic acid (EDTA), diethylenetriaminepentaacetic acid (DTMPA), alkyl- or alkenylsuccinic acid, soluble silicates or layered silicates (e.g. SKS-6 from Hoechst). The detergent may also be unbuilt, i.e. essentially free of detergent builder.

The detergent may comprise one or more polymers. Examples are carboxymethylcellulose (CMC), poly(vinylpyrrolidone) (PVP), polyethyleneglycol (PEG), poly(vinyl alcohol) (PVA), polycarboxylates such as polyacrylates, maleic/acrylic acid copolymers and lauryl methacrylate/acrylic acid copolymers.

The detergent may contain a bleaching system which may comprise a H_2O_2 source such as perborate or percarbonate which may be combined with a peracid-forming bleach activator such as tetraacetylethylenediamine (TAED) or nonanoyloxybenzenesulfonate (NOBS). Alternatively, the bleaching system may comprise peroxyacids of, e.g., the amide, imide, or sulfone type.

The enzymes of the detergent composition of the invention may be stabilized using conventional stabilizing agents, e.g. a polyol such as propylene glycol or glycerol, a sugar or sugar alcohol, lactic acid, boric acid, or a boric acid derivative such as, e.g., an aromatic borate ester, and the composition may be formulated as described in, e.g., WO 92/19709 and WO 92/19708.

The detergent may also contain other conventional detergent ingredients such as, e.g., fabric conditioners including clays, foam boosters, suds suppressors, anti-corrosion agents, soil-suspending agents, anti-soil-redeposition agents, dyes, bactericides, optical brighteners, or perfume.

The pH (measured in aqueous solution at use concentration) will usually be neutral or alkaline, e.g. in the range of 7-11.

Particular forms of detergent compositions within the scope of the invention include:

1) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

Linear alkylbenzenesulfonate (calculated as acid)	7 - 12%
Alcohol ethoxysulfate (e.g. C_{12-18} alcohol, 1-2 EO) or alkyl sulfate (e.g. C_{16-18})	1 - 4%
Alcohol ethoxylate (e.g. C_{14-15} alcohol, 7 EO)	5 - 9%
Sodium carbonate (as Na_2CO_3)	14 - 20%
Soluble silicate (as $Na_2O,2SiO_2$)	2 - 6%
Zeolite (as $NaAlSiO_4$)	15 - 22%
Sodium sulfate (as Na_2SO_4)	0 - 6%
Sodium citrate/citric acid (as $C_6H_5Na_3O_7/C_6H_8O_7$)	0 - 15%
Sodium perborate (as $NaBO_3.H_2O$)	11 - 18%
TAED	2 - 6%
Carboxymethylcellulose	0 - 2%
Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	0 - 3%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. suds suppressors, perfume, optical brightener, photobleach)	0 - 5%

2) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

Linear alkylbenzenesulfonate (calculated as acid)	6 - 11%
Alcohol ethoxysulfate (e.g. C_{12-18} alcohol, 1-2 EO or alkyl sulfate (e.g. C_{16-18})	1 - 3%
Alcohol ethoxylate (e.g. C_{14-15} alcohol, 7 EO)	5 - 9%
Sodium carbonate (as Na_2CO_3)	15 - 21%

Soluble silicate (as $Na_2O,2SiO_2$)	1 - 4%
Zeolite (as $NaAlSiO_4$)	24 - 34%
Sodium sulfate (as Na_2SO_4)	4 - 10%
Sodium citrate/citric acid (as $C_6H_5Na_3O_7/C_6H_8O_7$)	0 - 15%
Carboxymethylcellulose	0 - 2%
Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	1 - 6%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. suds suppressors, perfume)	0 - 5%

3) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

Linear alkylbenzenesulfonate (calculated as acid)	5 - 9%
Alcohol ethoxylate (e.g. C_{12-15} alcohol, 7 EO)	7 - 14%
Soap as fatty acid (e.g. C_{16-22} fatty acid)	1 - 3%
Sodium carbonate (as Na_2CO_3)	10 - 17%
Soluble silicate (as $Na_2O,2SiO_2$)	3 - 9%
Zeolite (as $NaAlSiO_4$)	23 - 33%
Sodium sulfate (as Na_2SO4)	0 - 4%
Sodium perborate (as $NaBO_3.H_2O$)	8 - 16%
TAED	2 - 8%
Phosphonate (e.g. EDTMPA)	0 - 1%
Carboxymethylcellulose	0 - 2%
Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	0 - 3%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. suds suppressors, perfume, optical brightener)	0 - 5%

4) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

Linear alkylbenzenesulfonate (calculated as acid)	8 - 12%
Alcohol ethoxylate (e.g. C_{12-15} alcohol, 7 EO)	10 - 25%
Sodium carbonate (as Na_2CO_3)	14 - 22%
Soluble silicate (as $Na_2O,2SiO_2$)	1 - 5%
Zeolite (as $NaAlSiO_4$)	25 - 35%
Sodium sulfate (as Na_2SO_4)	0 - 10%
Carboxymethylcellulose	0 - 2%
Polymers (e.g. maleic/acrylic acid copolymer, PVP, PEG)	1 - 3%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. suds suppressors, perfume)	0 - 5%

5) An aqueous liquid detergent composition comprising

Linear alkylbenzenesulfonate (calculated as acid)	15 - 21%
Alcohol ethoxylate (e.g. C_{12-15} alcohol, 7 EO or C_{12-15} alcohol, 5 EO)	12 - 18%
Soap as fatty acid (e.g. oleic acid)	3 - 13%
Alkenylsuccinic acid (C_{12-14})	0 - 13%
Aminoethanol	8 - 18%
Citric acid	2 - 8%
Phosphonate	0 - 3%
Polymers (e.g. PVP, PEG)	0 - 3%
Borate (as B_4O_7)	0 - 2%
Ethanol	0 - 3%
Propylene glycol	8 - 14%

Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. dispersants, suds suppressors, perfume, optical brightener)	0 - 5%

6) An aqueous structured liquid detergent composition comprising

Linear alkylbenzenesulfonate (calculated as acid)	15 - 21%
Alcohol ethoxylate (e.g. C_{12-15} alcohol, 7 EO, or C_{12-15} alcohol, 5 EO)	3 - 9%
Soap as fatty acid (e.g. oleic acid)	3 - 10%
Zeolite (as $NaAlSiO_4$)	14 - 22%
Potassium citrate	9 - 18%
Borate (as B_4O_7)	0 - 2%
Carboxymethylcellulose	0 - 2%
Polymers (e.g. PEG, PVP)	0 - 3%
Anchoring polymers such as, e.g., lauryl methacrylate/acrylic acid copolymer; molar ratio 25:1; MW 3800	0 - 3%
Glycerol	0 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. dispersants, suds suppressors, perfume, optical brighteners)	0 - 5%

7) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

Fatty alcohol sulfate	5 - 10%
Ethoxylated fatty acid monoethanolamide	3 - 9%
Soap as fatty acid	0 - 3%
Sodium carbonate (as Na_2CO_3)	5 - 10%

Soluble silicate (as $Na_2O,2SiO_2$)	1 - 4%
Zeolite (as $NaAlSiO_4$)	20 - 40%
Sodium sulfate (as Na_2SO_4)	2 - 8%
Sodium perborate (as $NaBO_3.H_2O$)	12 - 18%
TAED	2 - 7%
Polymers (e.g. maleic/acrylic acid copolymer, PEG)	1 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. optical brightener, suds suppressors, perfume)	0 - 5%

8) A detergent composition formulated as a granulate comprising

Linear alkylbenzenesulfonate (calculated as acid)	8 - 14%
Ethoxylated fatty acid monoethanolamide	5 - 11%
Soap as fatty acid	0 - 3%
Sodium carbonate (as Na_2CO_3)	4 - 10%
Soluble silicate (as $Na_2O,2SiO_2$)	1 - 4%
Zeolite (as $NaAlSiO_4$)	30 - 50%
Sodium sulfate (as Na_2SO_4)	3 - 11%
Sodium citrate (as $C_6H_5Na_3O_7$)	5 - 12%
Polymers (e.g. PVP, maleic/acrylic acid copolymer, PEG)	1 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. suds suppressors, perfume)	0 - 5%

9) A detergent composition formulated as a granulate comprising

Linear alkylbenzenesulfonate (calculated as acid)	6 - 12%
Nonionic surfactant	1 - 4%
Soap as fatty acid	2 - 6%

Sodium carbonate (as Na_2CO_3)	14 - 22%
Zeolite (as $NaAlSiO_4$)	18 - 32%
Sodium sulfate (as Na_2SO_4)	5 - 20%
Sodium citrate (as $C_6H_5Na_3O_7$)	3 - 8%
Sodium perborate (as $NaBO_3.H_2O$)	4 - 9%
Bleach activator (e.g. NOBS or TAED)	1 - 5%
Carboxymethylcellulose	0 - 2%
Polymers (e.g. polycarboxylate or PEG)	1 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. optical brightener, perfume)	0 - 5%

10) An aqueous liquid detergent composition comprising

Linear alkylbenzenesulfonate (calculated as acid)	15 - 23%
Alcohol ethoxysulfate (e.g. C_{12-15} alcohol, 2-3 EO)	8 - 15%
Alcohol ethoxylate (e.g. C_{12-15} alcohol, 7 EO, or C_{12-15} alcohol, 5 EO)	3 - 9%
Soap as fatty acid (e.g. lauric acid)	0 - 3%
Aminoethanol	1 - 5%
Sodium citrate	5 - 10%
Hydrotrope (e.g. sodium toluensulfonate)	2 - 6%
Borate (as B_4O_7)	0 - 2%
Carboxymethylcellulose	0 - 1%
Ethanol	1 - 3%
Propylene glycol	2 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%

Minor ingredients (e.g. polymers, dispersants, perfume, optical brighteners)	0 - 5%

11) An aqueous liquid detergent composition comprising

Linear alkylbenzenesulfonate (calculated as acid)	20 - 32%
Alcohol ethoxylate (e.g. C_{12-15} alcohol, 7 EO, or C_{12-15} alcohol, 5 EO)	6 - 12%
Aminoethanol	2 - 6%
Citric acid	8 - 14%
Borate (as B_4O_7)	1 - 3%
Polymer (e.g. maleic/acrylic acid copolymer, anchoring polymer such as, e.g., lauryl methacrylate/acrylic acid copolymer)	0 - 3%
Glycerol	3 - 8%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. hydrotropes, dispersants, perfume, optical brighteners)	0 - 5%

12) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

Anionic surfactant (linear alkylbenzenesulfonate, alkyl sulfate, alpha-olefinsulfonate, alpha-sulfo fatty acid methyl esters, alkanesulfonates, soap)	25 - 40%
Nonionic surfactant (e.g. alcohol ethoxylate)	1 - 10%
Sodium carbonate (as Na_2CO_3)	8 - 25%
Soluble silicates (as Na_2O, $2SiO_2$)	5 - 15%
Sodium sulfate (as Na_2SO_4)	0 - 5%
Zeolite (as $NaAlSiO_4$)	15 - 28%
Sodium perborate (as $NaBO_3.4H_2O$)	0 - 20%

Bleach activator (TAED or NOBS)	0 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. perfume, optical brighteners)	0 - 3%

13) Detergent formulations as described in 1) - 12) wherein all or part of the linear alkylbenzenesulfonate is replaced by (C_{12}-C_{18}) alkyl sulfate.

14) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

(C_{12}-C_{18}) alkyl sulfate	9 - 15%
Alcohol ethoxylate	3 - 6%
Polyhydroxy alkyl fatty acid amide	1 - 5%
Zeolite (as $NaAlSiO_4$)	10 - 20%
Layered disilicate (e.g. SK56 from Hoechst)	10 - 20%
Sodium carbonate (as Na_2CO_3)	3 - 12%
Soluble silicate (as $Na_2O,2SiO_2$)	0 - 6%
Sodium citrate	4 - 8%
Sodium percarbonate	13 - 22%
TAED	3 - 8%
Polymers (e.g. polycarboxylates and PVP=	0 - 5%
Enzymes (calculated as pure enzyme protein)	0.0001 - 0.1%
Minor ingredients (e.g. optical brightener, photo bleach, perfume, suds suppressors)	0 - 5%

15) A detergent composition formulated as a granulate having a bulk density of at least 600 g/l comprising

(C_{12}-C_{18}) alkyl sulfate	4 - 8%

Alcohol ethoxylate	11	- 15%
Soap	1	- 4%
Zeolite MAP or zeolite A	35	- 45%
Sodium carbonate (as Na_2CO_3)	2	- 8%
Soluble silicate (as $Na_2O,2SiO_2$)	0	- 4%
Sodium percarbonate	13	- 22%
TAED	1	- 8%
Carboxymethyl cellulose	0	- 3%
Polymers (e.g. polycarboxylates and PVP)	0	- 3%
Enzymes (calculated as pure enzyme protein)	0.0001	- 0.1%
Minor ingredients (e.g. optical brightener, phosphonate, perfume)	0	- 3%

16) Detergent formulations as described in 1) - 15) which contain a stabilized or encapsulated peracid, either as an additional component or as a substitute for already specified bleach systems.

17) Detergent compositions as described in 1), 3), 7), 9) and 12) wherein perborate is replaced by percarbonate.

18) Detergent compositions as described in 1), 3), 7), 9), 12), 14) and 15) which additionally contain a manganese catalyst. The manganese catalyst may, e.g., be one of the compounds described in "Efficient manganese catalysts for low-temperature bleaching", Nature 369, 1994, pp. 637-639.

19) Detergent composition formulated as a nonaqueous detergent liquid comprising a liquid nonionic surfactant such as, e.g., linear alkoxylated primary alcohol, a builder system (e.g. phosphate), enzyme and alkali. The detergent may also comprise anionic surfactant and/or a bleach system.

The protease variant of the invention may be incorporated in concentrations conventionally employed in detergents. It is at present contemplated that, in the detergent composition of the invention, the protease variant may be added in an amount corresponding to 0.00001-1 mg (calculated as pure enzyme protein) of protease variant per liter of wash liquor.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: NOVO NORDISK A/S
(B) STREET: Novo Alle
(C) CITY: Bagsvaerd
(E) COUNTRY: DENMARK
(F) POSTAL CODE (ZIP): DK-2880
(G) TELEPHONE: +45 44448888
(H) TELEFAX: +45 4449 3256
(I) TELEX: 37304

(ii) TITLE OF INVENTION: Trypsin-like Fusarium protease

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 998 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Fusarium oxysporum
(B) STRAIN: DSM 2672

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..998

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATCATCAACC ACTCTTCACT CTTCAACTCT CCTCTCTTGG ATATCTATCT CTTCACCATG 60

GTCAAGTTCG CTTCCGTCGT TGCACTTGTT GCTCCCCTGG CTGCTGCCGC TCCTCAGGAG 120

ATCCCCAACA TTGTTGGTGG CACTTCTGCC AGCGCTGGCG ACTTTCCCTT CATCGTGAGC 180

ATTAGCCGCA ACGGTGGCCC CTGGTGTGGA GGTTCTCTCC TCAACGCCAA CACCGTCTTG 240

ACTGCTGCCC ACTGCGTTTC CGGATACGCT CAGAGCGGTT TCCAGATTCG TGCTGGCAGT 300

CTGTCTCGCA CTTCTGGTGG TATTACCTCC TCGCTTTCCT CCGTCAGAGT TCACCCTAGC 360

TACAGCGGAA ACAACAACGA TCTTGCTATT CTGAAGCTCT CTACTTCCAT CCCCTCCGGC 420

GGAAACATCG GCTATGCTCG CCTGGCTGCT TCCGGCTCTG ACCCTGTCGC TGGATCTTCT 480

GCCACTGTTG CTGGCTGGGG CGCTACCTCT GAGGGCGGCA GCTCTACTCC CGTCAACCTT 540

CTGAAGGTTA CTGTCCCTAT CGTCTCTCGT GCTACCTGCC GAGCTCAGTA CGGCACCTCC 600

GCCATCACCA ACCAGATGTT CTGTGCTGGT GTTTCTTCCG GTGGCAAGGA CTCTTGCCAG 660

GGTGACAGCG GCGGCCCCAT CGTCGACAGC TCCAACACTC TTATCGGTGC TGTCTCTTGG 720

GGTAACGGAT GTGCCCGACC CAACTACTCT GGTGTCTATG CCAGCGTTGG TGCTCTCCGC 780

TCTTTCATTG ACACCTATGC TTAAATACCT TGTTGGAAGC GTCGAGATGT TCCTTGAATA 840

TTCTCTAGCT TGAGTCTTGG ATACGAAACC TGTTTGAGAA ATAGGTTTCA ACGAGTTAAG 900

AAGATATGAG TTGATTTCAG TTGGATCTTA GTCCTGGTTG CTCGTAATAG AGCAATCTAG 960

ATAGCCCAAA TTGAATATGA AATTTGATGA AAATATTC 998

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 248 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Fusarium oxysporum
(B) STRAIN: DSM 2672

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 25..248

(ix) FEATURE:

(A) NAME/KEY: Peptide
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /label= pre-propeptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Val Lys Phe Ala Ser Val Val Ala Leu Val Ala Pro Leu Ala Ala
                -20                 -15                 -10

Ala Ala Pro Gln Glu Ile Pro Asn Ile Val Gly Gly Thr Ser Ala Ser
            -5                  1                   5

Ala Gly Asp Phe Pro Phe Ile Val Ser Ile Ser Arg Asn Gly Gly Pro
    10                  15                  20

Trp Cys Gly Gly Ser Leu Leu Asn Ala Asn Thr Val Leu Thr Ala Ala
25                  30                  35                  40

His Cys Val Ser Gly Tyr Ala Gln Ser Gly Phe Gln Ile Arg Ala Gly
                45                  50                  55

Ser Leu Ser Arg Thr Ser Gly Gly Ile Thr Ser Ser Leu Ser Ser Val
            60                  65                  70

Arg Val His Pro Ser Tyr Ser Gly Asn Asn Asn Asp Leu Ala Ile Leu
        75                  80                  85

Lys Leu Ser Thr Ser Ile Pro Ser Gly Gly Asn Ile Gly Tyr Ala Arg
    90                  95                  100

Leu Ala Ala Ser Gly Ser Asp Pro Val Ala Gly Ser Ser Ala Thr Val
105                 110                 115                 120

Ala Gly Trp Gly Ala Thr Ser Glu Gly Gly Ser Ser Thr Pro Val Asn
                125                 130                 135
```

Leu Leu Lys Val Thr Val Pro Ile Val Ser Arg Ala Thr Cys Arg Ala
140 145 150

Gln Tyr Gly Thr Ser Ala Ile Thr Asn Gln Met Phe Cys Ala Gly Val
155 160 165

Ser Ser Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Ile
170 175 180

Val Asp Ser Ser Asn Thr Leu Ile Gly Ala Val Ser Trp Gly Asn Gly
185 190 195 200

Cys Ala Arg Pro Asn Tyr Ser Gly Val Tyr Ala Ser Val Gly Ala Leu
205 210 215

Arg Ser Phe Ile Asp Thr Tyr Ala
220

CLAIMS

1. A variant of a parent trypsin-like Fusarium protease comprising the amino acid sequence shown in the appended Sequence Listing ID No. 2, which

i) reacts with an antibody raised against or reactive with at least one epitope of the parent trypsin-like protease,

ii) is at least 60% homologous with the amino acid sequence of the parent trypsin-like protease, and/or

iii) is encoded by a DNA sequence which hybridizes with an oligonucleotide probe hybridizing with the trypsin-like F. oxysporum protease comprising the amino acid sequence shown in SEQ ID No. 2.

2. A trypsin-like Fusarium protease variant according to claim 1, in which a naturally occurring amino acid residue (other than proline) of the amino acid sequence of the parent trypsin-like protease has been substituted with a proline residue at one or more positions, at which positions(s) the dihedral angles ϕ (phi) and ψ (psi) constitute values within the intervals [$-90° < \phi < -40°$ and $-180° < \psi < 180°$], preferably within the intervals [$-90° < \phi < -40°$ and $120° < \psi < 180°$] or [$-90° < \phi < -40°$ and $-50° < \psi < 10°$], and which position(s) is/are not located in regions in which the protease is characterized by possessing α-helical or ß-sheet structure.

3. The variant according to claim 2, in which the amino acid residue(s) to be substituted with proline is a hydrophilic or a small hydrophobic amino acid residue, in particular one selected from the amino acid residues A, D, E, K, G, Q, R, S, T, N and V.

4. The variant according to claim 2, in which one or more of the amino acid residues occupying the positions 21, 24, 30, 46, 49, 71, 90, 111, 114, 121, 124, 125, 126, 132, 135, 146, 148, 174, 175, 176, 178, 185b, 185c, 202, 208, and/or 209

of the parent trypsin-like protease (using the amino acid numbering defined in Table 1) has/have been substituted with proline residues.

5. A trypsin-like Fusarium protease variant according to claim 1, in which an amino acid residue different from cysteine of the amino acid sequence of the parent trypsin-like protease has been substituted with a cysteine residue in such a manner that the introduced cysteine residue together with another cysteine residue present in the parent protease or introduced therein form a disulphide bridge.

6. The variant according to claim 5, in which the amino acid residues occupying the positions 22+157, 129+232 and/or 136+201 of the parent trypsin-like protease (using the amino acid numbering defined in Table 1) have been substituted with cysteine residues.

7. The variant according to claim 6 comprising the following substitutions:
A22C+D26S+K157C, or
A136C+D201C.

8. A trypsin-like Fusarium protease variant according to claim 1, in which either or both residues of any of the Asn-Gly sequence appearing in the amino acid sequence of the parent trypsin-like protease is/are deleted or substituted with a residue of a different amino acid.

9. The variant according to claim 8, in which the Asn and/or Gly residue is substituted with a residue of an amino acid selected from the group consisting of A, Q, S, P, T and Y.

10. The variant according to claim 8, in which any of the Asn or Gly residues of the Asn-Gly occupying positions

36+38 and/or 217+219 of the parent trypsin-like protease (using the amino acid numbering defined in Table 1) has/have been deleted or substituted with a residue of an amino acid selected from the group consisting of A, Q, S, P, T and Y.

11. The variant according to claim 10 comprising one or more of the following substitutions:
N36S, G38S, N217Y,S, G119S

12. A trypsin-like Fusarium protease variant, in which one or more amino acid residues present in a loop structure of the parent trypsin-like protease suceptible to cleavage by a proteolytic enzyme is/are deleted or replaced with one or more amino acid residues so as to obtain a modified loop structure having an improved proteolytic stability.

13. The variant according to claim 12, in which the loop structure is suceptible to cleavage by a subtilisin or trypsin.

14. The variant according to claim 12, in which one or more amino acid substitutions has/have been made so that the modified loop structure comprises at least one of the amino acid segments GAAG, GARG, YPGS, YPRS, HNRG, YTGN, ISSE, NNAG, SFIN, DQNG, ASFS, SRGV, LDTG, YYAA, INDI, WYFG, SIEN, GSTY, DSTN, PDLR, LDTG, GNRY, SGVM, RYPS, NGLV, SFSI, LGSP, RASF, VPWG, PDLN, SFVP, PDYR, PRLP, TVLP, IGTC, TGGT, TNKL, VGDV, IGVL, GSTY, RYAN, PNIP, or TLVP.

15. The variant according to claim 12, in which loop II of the trypsin-like protease comprising the peptide sequence SRNGGP is substituted with loop II of the trypsin 2ptn isolated from Bos taurus comprising the peptide sequence NSGYH, as follows:

S34N+R35S+N36*+G39Y+P40H, and/or

loop IV of the trypsin-like protease comprising the peptide sequence VSGYAQSGF is substituted with loop IV of the trypsin 2ptn isolated from Bos taurus comprising the peptide sequence YKSGI, as follows:

V59Y+S59a*+G59b*+Y59c*+A60K+Q61S+S62*+F64I, and/or

loop IV of the trypsin-like protease comprising the peptide sequence VSGYAQSGF is substituted with loop IV of the trypsin 1trm isolated from Rattus rattus comprising the peptide sequence YKSRI, as follows:

V59Y+S59a*+G59b*+Y59c*+A60K+Q61S+S62*+G63R+F64I, and/or

loop VI of the trypsin-like protease comprising the peptide sequence HPSYSGN is substituted with loop VI of the trypsin 1trm isolated from Rattus rattus comprising the peptide sequence HPNFDRKTL, as follows:

S93N+F94Y+S95D+*96R+*97K+G98T+N99L.

16. A trypsin-like Fusarium protease variant according to claim 1, in which an autoproteolysis site of the parent trypsin-like protease has been removed.

17. The variant according to claim 16, in which an Arg and/or Lys residue of the parent trypsin-like protease has been substituted with a non or less positively charged amino acid residue, or in which an Arg residue has been substituted with a Lys residue.

18. The variant according to claim 17, in which the non or less positively charged amino acid residue is selected from the group consisting of K, S, V, P, E, D, N, Q, A and G.

19. The variant according to claim 16, in which the Lys and/or Arg residue at the following positions have been substituted:

35, 67, 73, 89, 107, 122, 157, 165, 169, 188, 222, 235, using the structural sequence numbering defined in Table 1.

20. The variant according to claim 19, comprising one or more of the following substitutions:

R35K,S
R122K,V,P
R169K,V
R122K,V,P+A124P
R122V,P+A124P+T208V

21. A trypsin-like Fusarium protease variant according to claim 1, in which one or more amino acid residues susceptible to oxidation has/have been replaced with another amino acid residue less susceptible to oxidation.

22. The variant according to claim 21, in which the amino acid residue less susceptible to oxidation is selected from the group consisting of A, E, N, Q, I, L, S and K.

23. The variant according to claim 21 comprising one or more of the following substitutions:

M180N,Q,E,K, using the sequence numbering of Table 1.

24. A trypsin-like Fusarium protease variant according to claim 1, in which a N-glycosylation site has been introduced at an amino acid residue X located in a loop structure on the surface of the protein subject to proteolysis by changing the sequence segment X-Y-Z to Asn-Y1-Ser or Asn-Y1-Thr, pro-

vided that Y1 is different from Gly, so as to confer to the variant an improved proteolysis resistance.

25. The variant according to claim 24, in which X is the amino acid residue located in the position 17, 18, 20, 21, 23, 24, 25, 34, 35, 36, 38, 39, 48, 49, 50, 59a, 59b, 59c, 60, 61, 62, 63, 69, 70, 71, 72, 73, 74, 78, 79, 80, 81, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 98, 99, 109, 110, 111, 113, 114, 115, 116, 117, 118, 121, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 163, 164, 165, 173, 174, 175, 176, 177a, 178, 184, 185, 185b, 187, 188, 201, 202, 203, 207, 215, 216, 217, 221, 222, 223, 224, 225, 233, 234, 235 or 236 of the parent trypsin-like protease using the amino acid numbering defined in Table 1.

26. The variant according to claim 24, in which the residue Y is Lys or Arg.

27. The variant according to claim 26, comprising one or more of the following substitutions:

S34N+R35R,K+N36S,T
N36N+G38S+G39S,T
S72N+R73R,K+T74T,S
V88N+R89R,K+V90S,T
A121N+R122R,K+L123S,T
L123N+A224A+A225S,T
S164N+R165R,K+A166S,T

28. A trypsin-like Fusarium protease variant according to claim 1, in which the net electrostatic charge of the parent trypsin-like protease has been changed by deleting or substituting one or more negatively charged amino acid residues by neutral or positively charged amino acid residue(s), and/or by substituting one or more neutral amino acid residues by

positively or negatively charged amino acid residue(s), and/or by deleting or substituting one or more positively charged amino acid residue(s) by neutral or negatively charged amino acid residue(s).

29. A variant according to claim 28, which has a lower pI as compared to the pI of its parent protease.

30. A variant according to claim 28, which has a higher pI as compared to that of its parent protease.

31. The variant according to claim 30, in which the amino acid residue occupying one or more of the positions 17, 18, 20, 21, 23, 24, 25, 26, 27, 34, 35, 36, 38, 39, 40, 41, 48, 49, 50, 59a, 59b, 59c, 60, 61, 62, 63, 65, 67, 69, 70, 71, 72, 73, 74, 75, 76, 78, 79, 80, 81, 82, 83, 86, 87, 88, 89, 90, 91, 93, 94, 95, 98, 99, 100, 101, 109, 110, 111, 113, 114, 115, 116, 117, 118, 119, 120, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 137, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 156, 157, 159, 161, 163, 164, 165, 166, 167, 169, 170, 171, 173, 174, 175, 176, 177a, 178, 179, 184, 185, 185b, 185c, 187, 188, 189, 192, 201, 202, 207, 208, 210, 215, 216, 217, 219, 221, 222, 223, 223a, 224, 225, 226, 233, 234, 235, 236, 237, 239, 240, 241 or 242 of the parent trypsin-like protease (using the amino acid numbering defined in Table 1) have been deleted or substituted with a neutral, negative or positive hydrophilic amino acid, preferably selected from E, D, K, R, S, A, P or T.

32. The variant according to claim 28, in which one or more amino acids around the active site has been substituted with an other amino acid sequence so as to obtain a change in the polarity around the active site.

33. The variant according to claim 32, in which the amino acid sequence occupying the position 214 of the parent

trypsin-like protease (using the amino acid numbering defined in Table 1) has been substituted with another amino acid.

34. The variant according to claim 33, in which the following substitution has been made:
S214A.

35. A variant of a parent trypsin-like protease, in which at least one of the loop structures II, IV, VI, IX, X and XIII has been substituted with a loop structure, which

a) is at least 60% homologous to the corresponding loop structure in the trypsin-like Fusarium protease comprising the amino acid sequence shown in SEQ ID No. 2, and/or

b) reacts with an antibody raised against the corresponding loop structure of the trypsin-like Fusarium protease comprising the amino acid sequence SEQ ID No. 2 and which recognizes said trypsin-like Fusarium protease.

36. The variant according to claim 35, in which the parent trypsin-like protease is derived from a microorganism, a mammal, a vertebrate or an insect.

37. The variant according to claim 36, in which the parent trypsin-like protease is derived from any of the mammalian trypsin-like protease listed in Table 2 herein, or is a trypsin-like protease derivable from the bacterial species Streptomyces, in particular from S. griseus.

38. A DNA construct comprising a DNA sequence encoding a trypsin-like protease variant according to any of claims 1-37.

39. A recombinant expression vector which carries a DNA construct according to claim 38.

40. A cell which is transformed with a DNA construct according to claim 38 or a vector according to claim 39.

41. A cell according to claim 40, which is a microbial cell.

42. A cell according to claim 41 which is a bacterial cell or a fungal cell.

43. A cell according to claim 42, in which the bacterial cell is a cell of a gram-positive bacterium, e.g. of the genus Bacillus or Streptomyces or a cell of a gram-negative bacterium, e.g. of the genus Escherichia, and the fungal cell is a yeast cell, e.g. of the genus Saccharomyces, or a cell of a filamentous fungus, e.g. of the genus Aspergillus or Fusarium.

44. A method of producing a variant of a trypsin-like protease according to any of claims 1-37, wherein a cell according to any of claims 40-43 is cultured under conditions conducive to the production of the variant, and the variant is subsequently recovered from the culture.

45. A detergent additive comprising a trypsin-like protease variant according to any of claims 1-37, optionally in the form of a non-dusting granulate, stabilised liquid or protected enzyme.

46. A detergent additive according to claim 45 which contains 0.02-200 mg of enzyme protein/g of the additive.

47. A detergent additive according to claim 45, which additionally comprises another enzyme such as a lipase, a protease, amylase, peroxidase and/or cellulase.

sp387

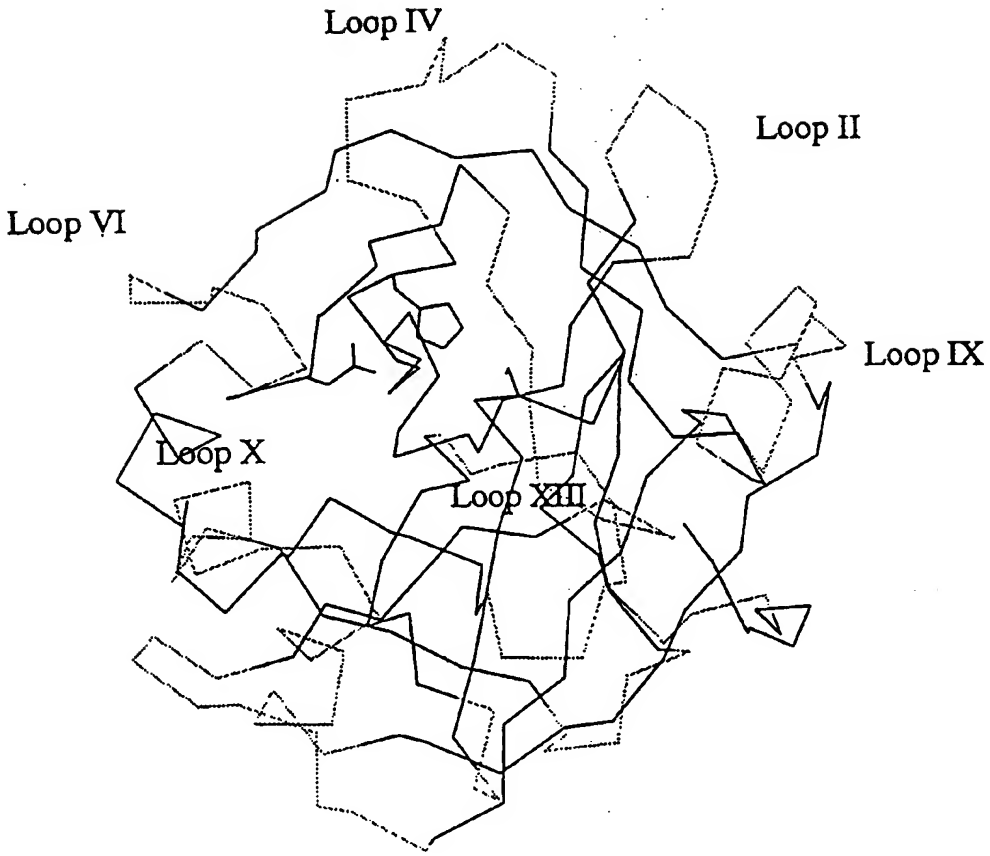


Fig. 1

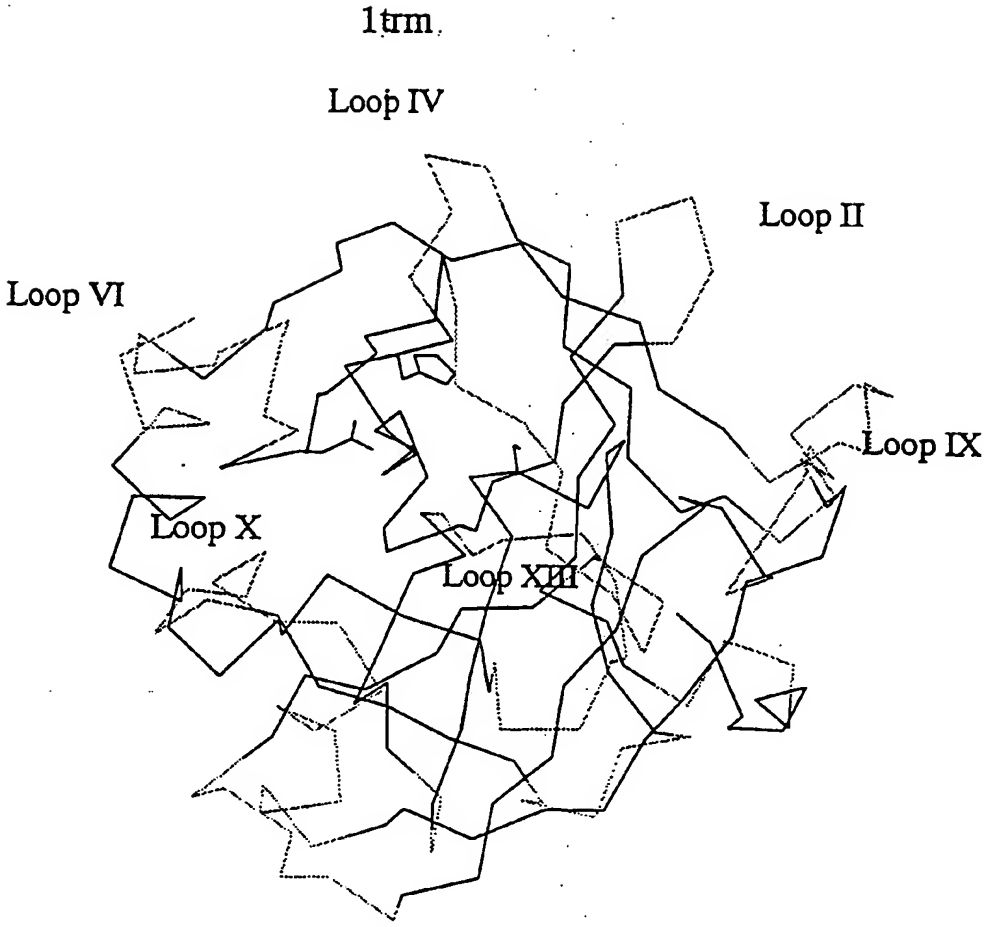


Fig. 2

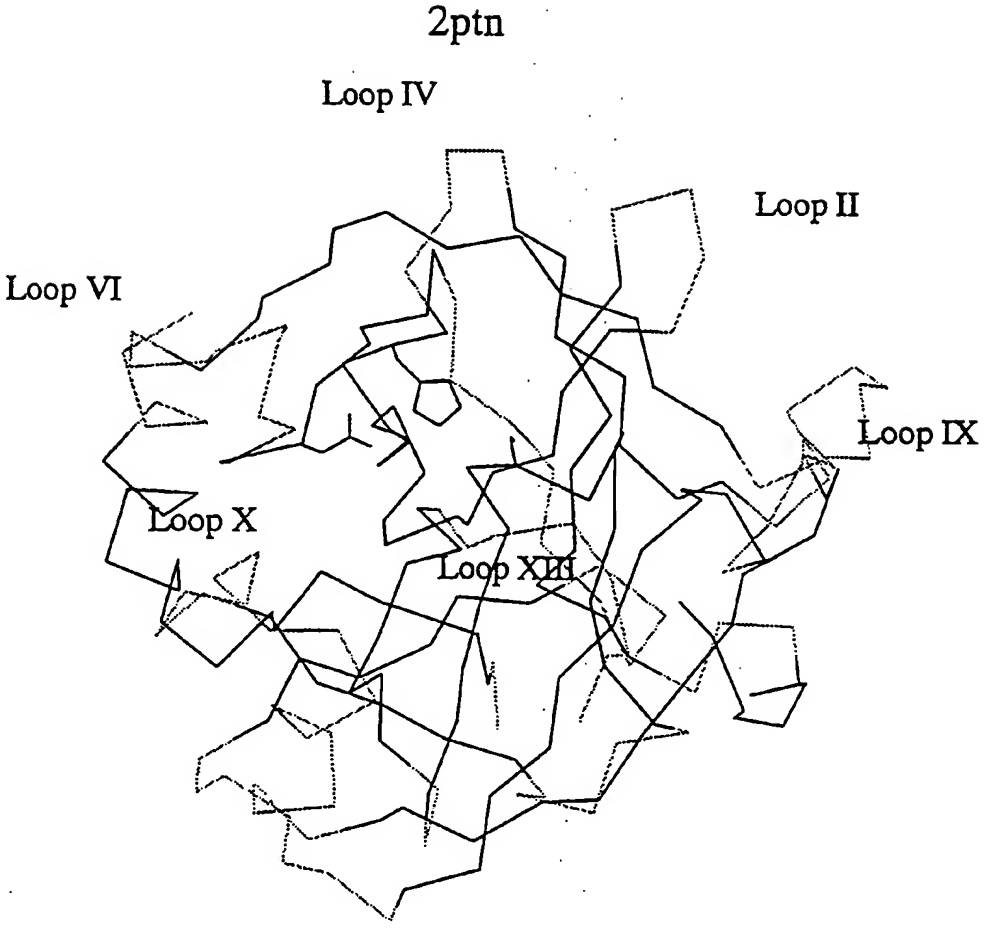


Fig. 3

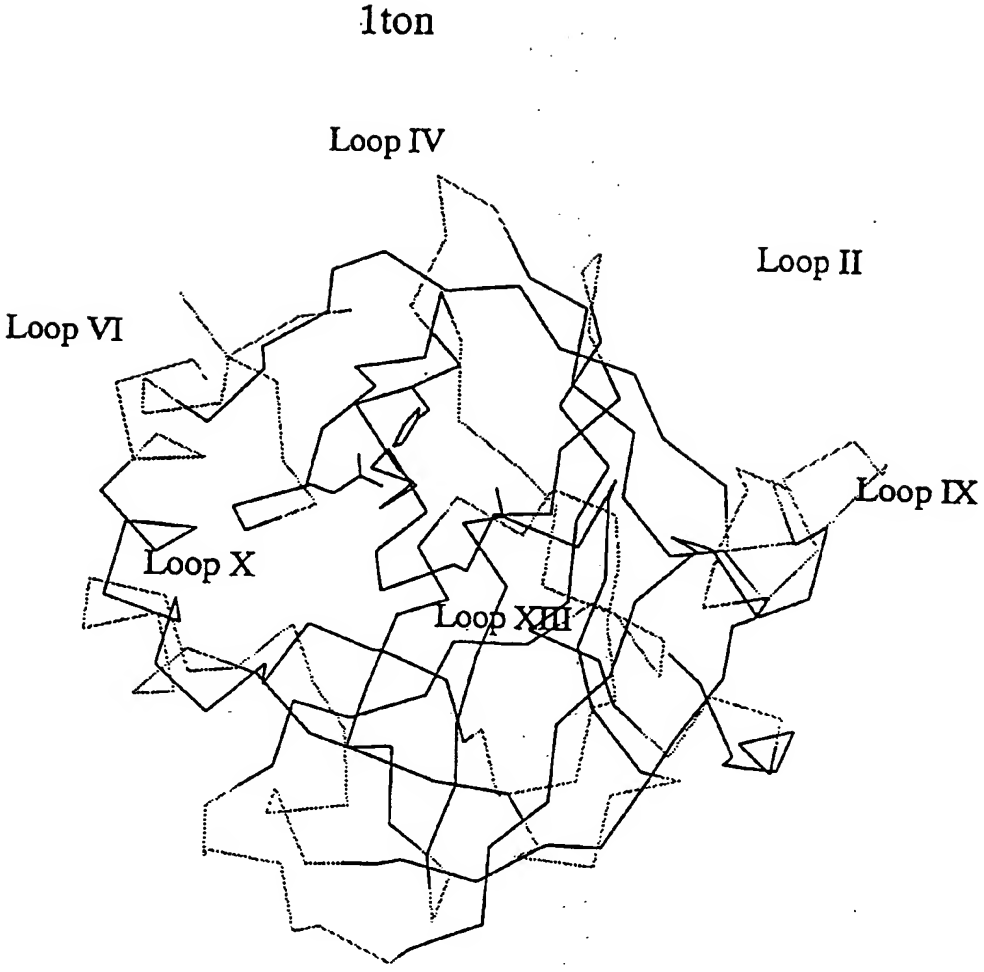


Fig. 4

2pka

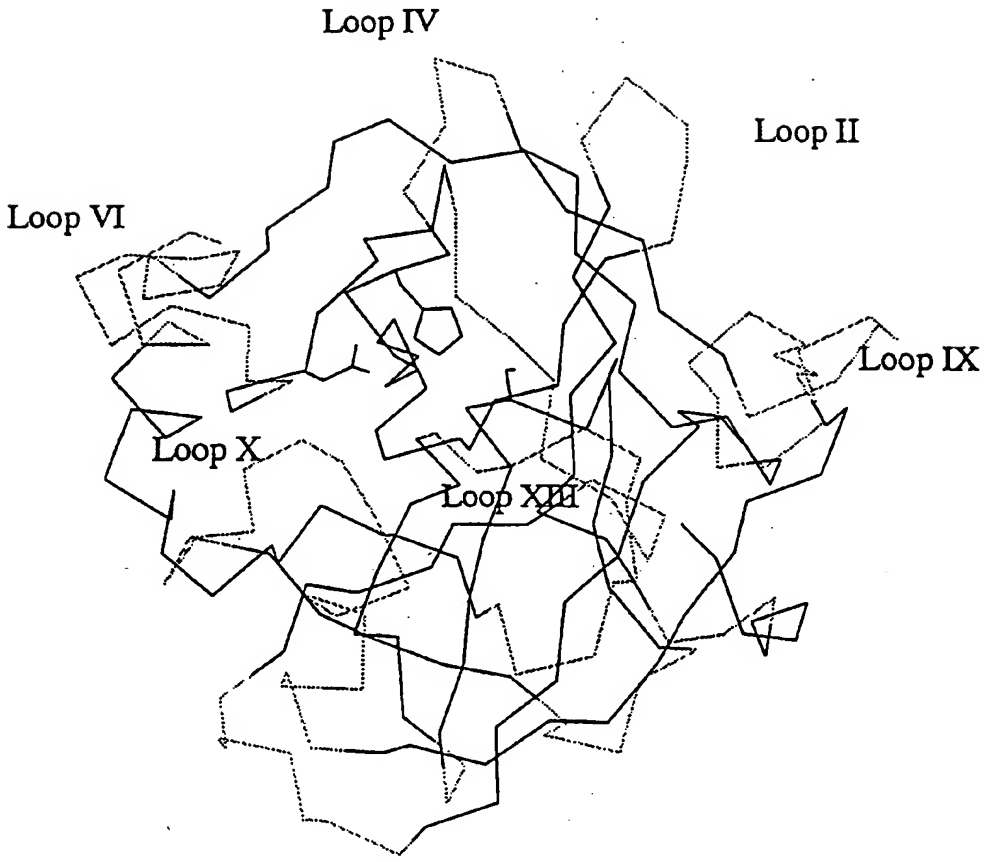


Fig. 5

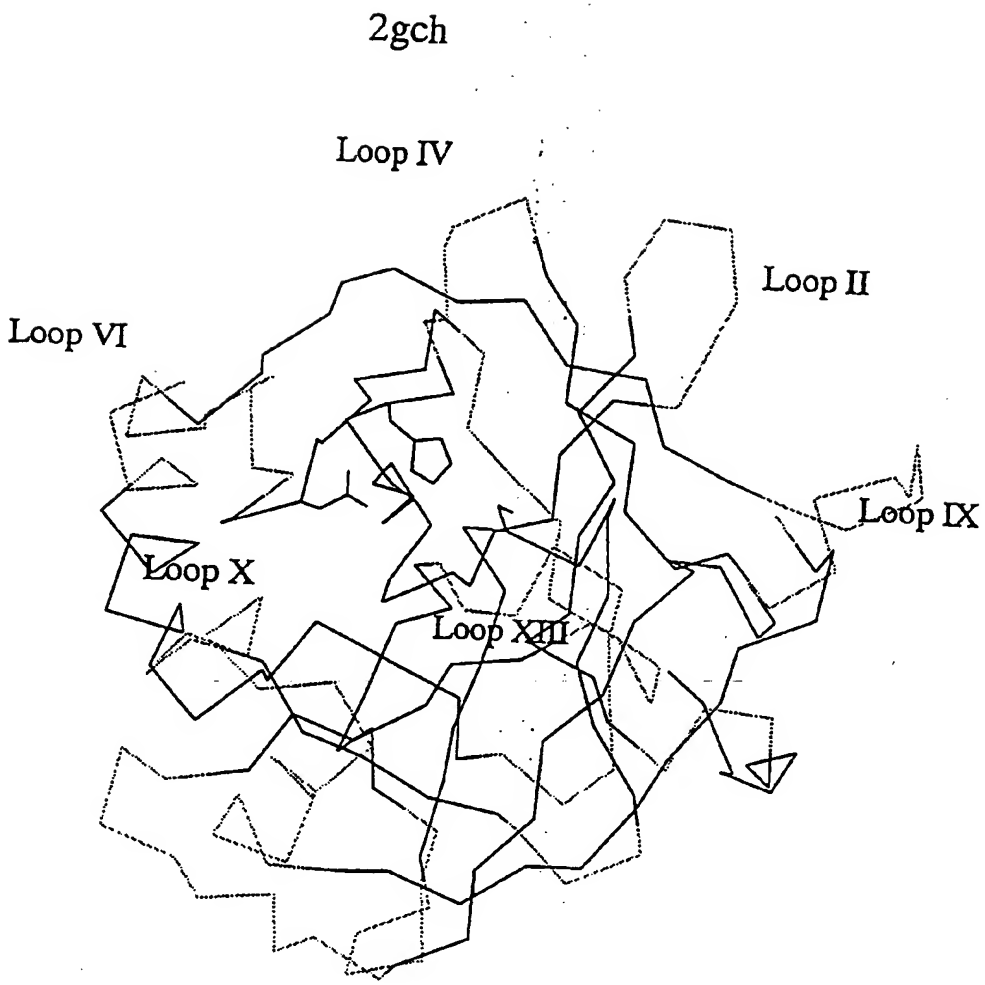


Fig. 6

3est

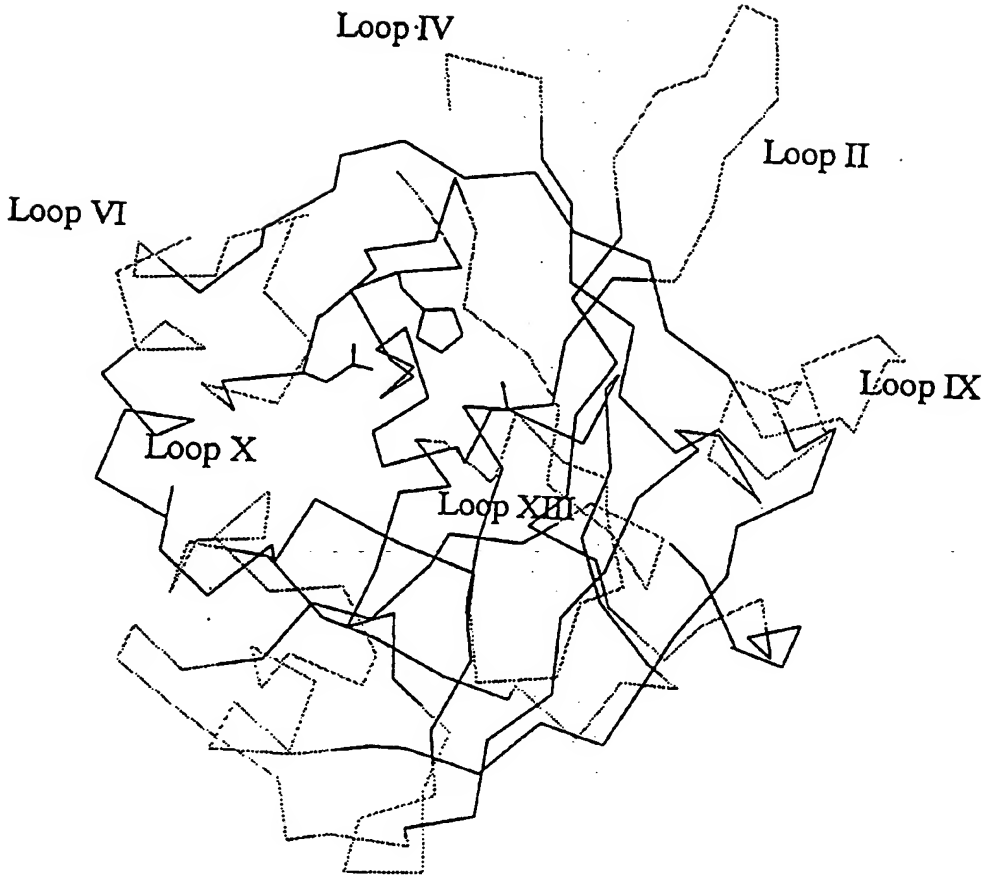


Fig. 7

8/10

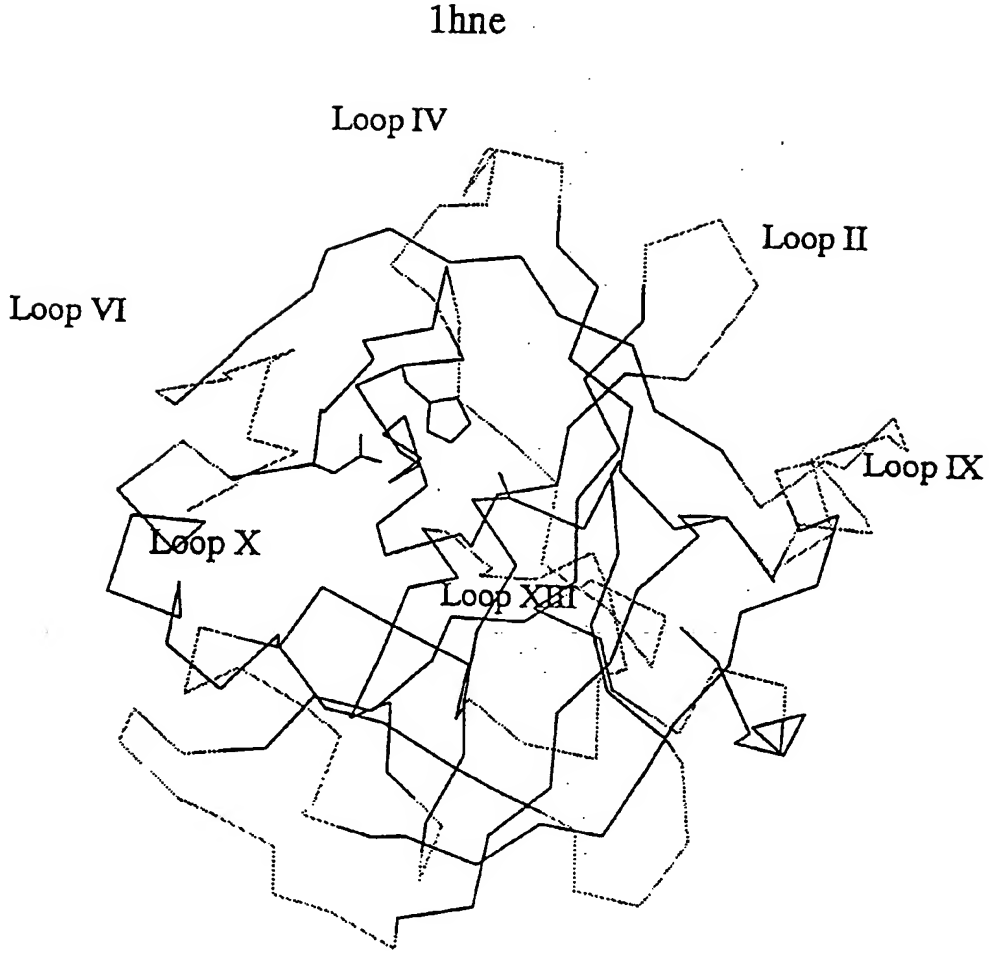


Fig. 8

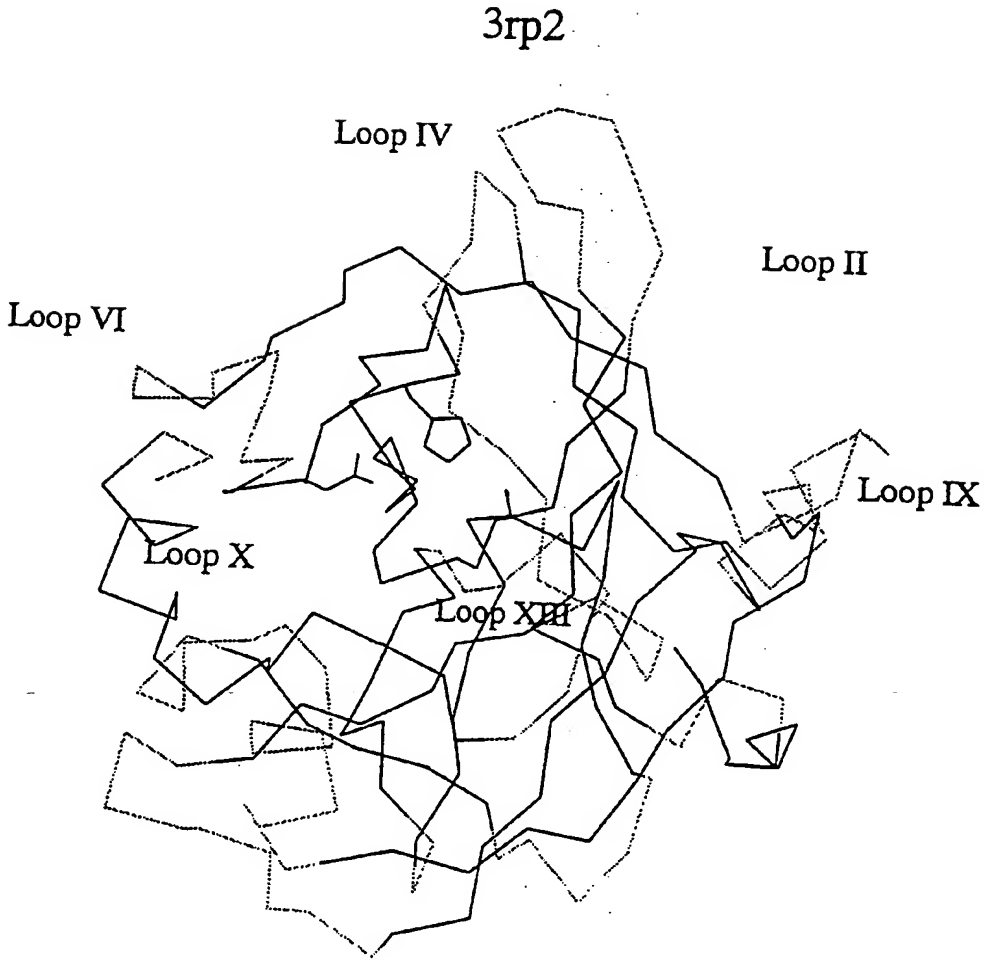


Fig. 9

10/10

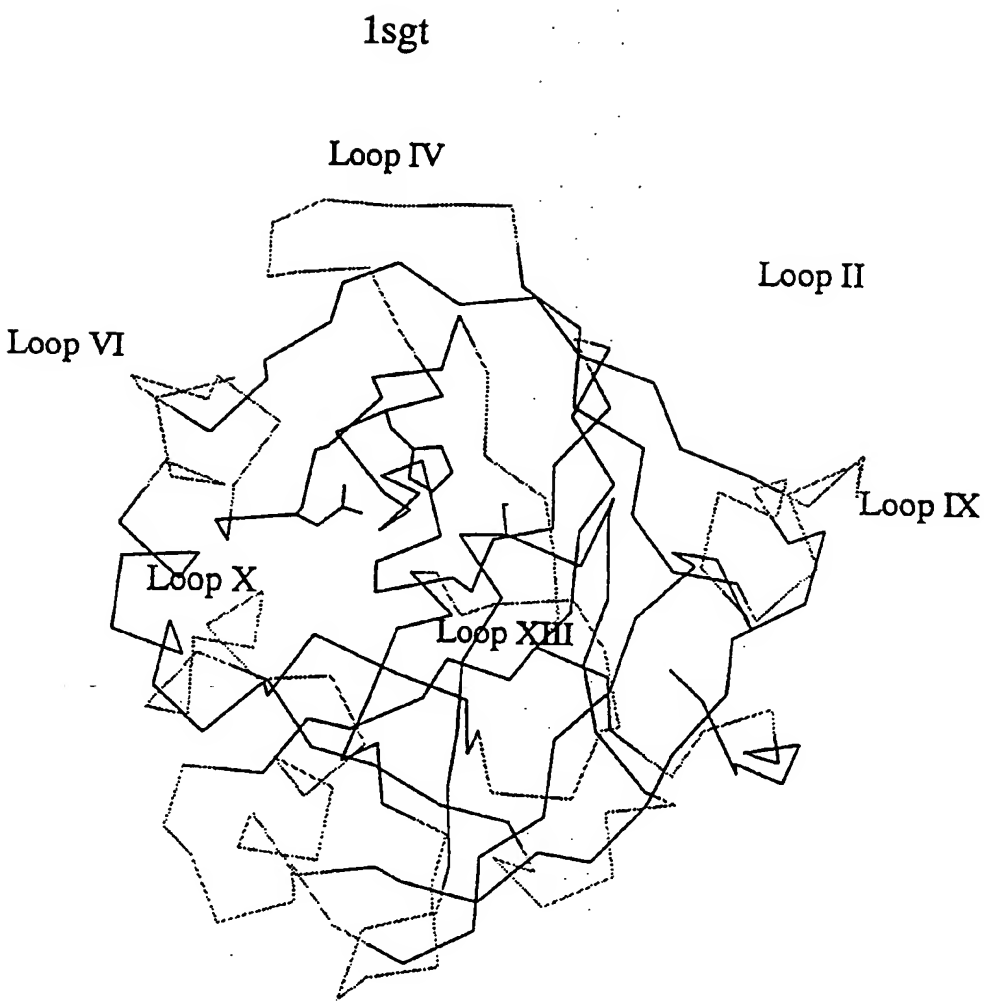


Fig. 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 95/00180

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12N 9/04, C12N 15/53
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 8906270 A1 (NOVO INDUSTRI A/S), 13 July 1989 (13.07.89), the whole document ---	1-4,38-47
X	WO 9219729 A1 (NOVO NORDISK A/S), 12 November 1992 (12.11.92)	2-4
A	--- ---	38-47

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 August 1995	06 -09- 1995
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Patrick Andersson Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 95/00180

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see extra sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 95/00180

Box III TEXT OF THE ABSTRACT (Continuation of item 5 of the first sheet)

According to rule 13.2, an international application shall relate to one invention only or a group of inventions linked by one or more of the same or corresponding "special technical features", i.e. features that define a contribution which each of the inventions makes over prior art.

Such a link between all the subject of claims 1-47 would be a variant of a parent trypsin-like Fusarium protease comprising the amino acid sequence shown in Sequence Listing ID No. 2. This *a priori* allegation however is not acceptable due to the state of the art as revealed in the attached search report i.e. WO 89/06270. This document shows the enzyme encoded by SEQ ID No.2., which enzyme is sequenced and cloned by methods well known in the art (description page 35, line 26-29). Accordingly, the following inventions were found:

*Invention1., claims 1-2 completely and 3-4, 44-47 partially : A variant of trypsin-like Fusarium protease comprising the amino acid sequence in which a naturally occurring amino acid residue, other than proline, has been substituted at positions where the dihedral angles phi and psi constitute values within the intervals [-90<phi<-40 degree and -180<psi<180 degree] but not being located in regions in which the protease is characterised by possessing alpha-helical or beta-sheet structure. The substituted amino acid residue being alanine.

*Invention 2-11, claims 3-4 and 44-47 partially: A variant of trypsin-like Fusarium protease comprising the amino acid sequence in which a naturally occurring amino acid residue, other than proline, has been substituted at positions where the dihedral angles phi and psi constitute values within the intervals [-90<phi<-40 degree and -180<psi<180 degree] but not being located in regions in which the protease is characterised by possessing alpha-helical or beta-sheet structure. The substituted amino acid residue being aspartic acid, glutamic acid, lysine, glycine, glutamine, arginine, serine, threonine, aspargine or valine respectively.

*A unifying "special technical feature" between inventions1-11 could have been to change amino acids at positions with certain defined dihedral angles, this allegation is not acceptable due to the state of the art revealed in WO92/19729.

Invention 12, claims 5-7 completely and 38-47 partially: A variant of trypsin-like Fusarium protease in which a naturally occurring amino acid residue, other than cystein, has been substituted with cystein.

Invention 13, claims 8-11 completely and 38-47 partially : A variant of trypsin-like Fusarium protease in which either one or both residues of any of the Asn-Gly sequence appearing in the parent protease is deleted or substituted with a different amino acid residue.

Box III TEXT OF THE ABSTRACT (Continuation of Item 5 of the first sheet)

Invention 14, claims 12-15 completely and 44-47 partially : A variant of trypsin-like Fusarium protease in which one or more amino acid residues present in a loop structure of the parent protease susceptible to cleavage by a proteolytic enzyme is deleted or replaced to obtain an enzyme having improved proteolytic stability.

Invention 15, claims 16-20 completely and 38-47 partially: A variant of trypsin-like Fusarium protease comprising SEQ ID No. 2 in which an autoproteolytic site of the parent protease have been removed.

Invention 16, claims 21-23 completely and 38-47 partially: A variant of trypsin-like Fusarium protease comprising SEQ ID No. 2 in which one or more amino acid residues of the parent protease susceptible to oxidation has been replaced with another amino acid residue less susceptible to oxidation.

Invention 17, claims 24-27 completely and 38-47 partially: A variant of trypsin-like Fusarium protease comprising SEQ ID No. 2 in which a N-glycosylation site has been introduced at an amino acid residue X located in a loop structure on the surface of the protein subject to proteolysis by changing the sequence segment X-Y-Z to Asn-Y1-Ser or Asn-Y1-Thr, provided that Y1 is different from Gly, so as to confer to the variant an improved proteolysis resistance.

Invention 18, claims 28-34 completely and 38-47 partially: A variant of trypsin-like Fusarium protease comprising SEQ ID No. 2 in which the net electrostatic charge of the parent enzyme has been changed by deleting or substituting one or more amino acid residue with an amino acid with different charge.

Invention 19, claims 35-43 completely and 38-47 partially: A variant of trypsin-like Fusarium protease in which at least one of the loop structures II, IV, VI, IX, X and XIII has been substituted with a loop structure homologous to SEQ ID No. 2.

The search has been restricted to inventions 1-11.
The division of inventions 12-19 are only general and may at a further search be found to be even a larger number of inventions.

INTERNATIONAL SEARCH REPORT

31/07/95

International application No.
PCT/DK 95/00180

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A1- 8906270	13/07/89	DE-U- 6890098	16/04/92
		EP-A,B- 0394352	31/10/90
		EP-A- 0471265	19/02/92
		JP-T- 3502203	23/05/91
		US-A- 5288627	22/02/94
WO-A1- 9219729	12/11/92	EP-A- 0583339	23/02/94
		JP-T- 6506832	04/08/94